2019

Shaping the perceptual representation of observed human action through prediction.

McDonough, Katrina Louise

Copyright statement

UNIVERSITY OF
PLYMOUTH

SHAPING THE PERCEPTUAL REPRESENTATION OF OBSERVED HUMAN ACTION THROUGH PREDICTION

by

KATRINA LOUISE MCDONOUGH

A thesis submitted to the University of Plymouth

in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

June 2019

Acknowledgements

First and foremost, I would like to express huge gratitude to my supervisor, Patric Bach, whose unwavering support, wisdom, and kindness have known no bounds throughout my work with him. He invested in me as an undergraduate and has inspired me in every step of the way with his intelligence and passion for research. His never-ending encouragement and optimism, together with his cheerful nature and good humour, has made this journey an enjoyable and stress-free experience.

I give special thanks to Matthew Hudson, who showed constant support and encouragement during my undergraduate and throughout the PhD process. His unique ideas and stimulating discussions have been an inspiration and have helped me to develop my own creative thinking and research ideas.

I owe an enormous debt to all the lovely fellow PhD students, past and present, and to all my Plymouth friends who have always shown great support, comfort and friendship and, above all, much laughter. They have always been there in times of need and definitely know how to bring the fun.

I would like to thank my family for their love and support. Their continued praise and encouragement has given me the motivation to carry on through the hardest times.

Above all, I owe the greatest debt to my boyfriend, John. I am eternally grateful for his endless love, support, patience and undoubting belief in me that has been a constant reminder of how lucky I am. His rational thinking and words of wisdom have saved me from despair countless times, and his continued faith in me has given me the confidence to believe in myself. He is, and always will be, my hero.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

This study was financed with the aid of a studentship from the University of Plymouth. In addition, a bursary was awarded by University of Plymouth to attend the Visceral Minds summer school at Bangor University in 2016, a second bursary was awarded in 2017 to support participant recruitment and a third bursary was awarded to attend the Concepts, Actions and Objects Conference in 2017.

The research in thesis has been published in scientific journals and has been presented as talks and posters at various conferences (listed below).

Publications:

Chapter Three and Four:

Hudson*, M., **McDonough*, K. L**., Edwards, R., & Bach*, P. (2018). Perceptual Teleology: Expectations of Action Efficiency Bias Social Perception. *Proceedings of the Royal Society: Biological sciences*. *285*(1884), 20180638. doi:10.1098/rspb.2018.0638 (***authors contributed equally**)

Chapter Five:

McDonough, K. L., Hudson, M., & Bach, P. (2019). Cues to intention bias action perception toward the most efficient trajectory. *Scientific Reports*, *9*(1), 6472. doi:10.1038/s41598-019-42204-y

Chapter Six:

McDonough, K. L., Costantini, M., Hudson, M., & Bach, P. (*under review*). Affordance matching predictively shapes the perceptual representation of others' ongoing actions. Preprint available: https://psyarxiv.com/mbea2

Conference Talks:

McDonough, K. L., Hudson, M., Bach, P. (2019). Expectations of Action Bias Social Perception. 11th Annual School of Psychology Conference, University of Plymouth, UK.

McDonough, K. L., Hudson, M., Bach, P. (2018). Expectations of Action Bias Social Perception. 10th Annual School of Psychology Conference, University of Plymouth, UK.

McDonough, K. L., Hudson, M., Bach, P. (2018). Expectations of Action Bias Social Perception. MeeTo 2018, University of Turin, Italy.

McDonough, K. L., Hudson, M. & Bach, P. (2016). Missing in Action: Prior expectations bias perception of vanishing goal directed hand actions. 8th Annual School of Psychology Conference, University of Plymouth, UK.

Conference Posters:

McDonough, K. L., Costantini, M., Hudson, M., & Bach, P. (2018). Getting to grips with action understanding: Affordance matching predictively shapes the perceptual representation of others' actions. Future of Social Cognition conference, University of East Anglia, UK.

McDonough, K. L., Costantini, M., Hudson, M., & Bach, P. (2018). Getting to grips with action understanding: Affordance matching predictively shapes the perceptual representation of others' actions. MeeTo 2018, University of Turin, Italy.

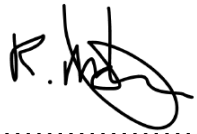
McDonough, K. L., Costantini, M., Hudson, M., & Bach, P. (2017). Getting to grips Affordance understanding: Affordance matching predictively shapes the perceptual representation of others' actions. British Association for Cognitive Neuroscience Annual Meeting, University of Plymouth, UK.

McDonough, K. L., Hudson, M. & Bach, P. (2017). Seeing is Believing? Prior knowledge of others' beliefs biases perception of their actions. British Association for Cognitive Neuroscience Annual Meeting, University of Plymouth, UK.

McDonough, K. L., Hudson, M. & Bach, P. (2017). Seeing is Believing? Prior knowledge of others' beliefs biases perception of their actions. Concepts, Actions and Objects: Functional and Neural Perspectives Workshop. University of Trento, Italy.

McDonough, K. L., Hudson, M. & Bach, P. (2016). Assumptions of rational action bias the perception of others' actions, 8th Annual School of Psychology Conference, University of Plymouth, UK.

Word count of main body of thesis: 44,452 words.



Signed …….

Date ……03/06/2019…….

Shaping the perceptual representation of observed human action through prediction.

Katrina Louise McDonough

Abstract

Understanding the actions of others is crucial for all social interactions. Despite a dynamic and complicated social world, humans can derive the goals, attitudes and beliefs that drive others' actions, imbuing them with meaning and understanding. While such abilities were traditionally accounted for by a direct matching of observed actions to actions within the observer's motor system, contemporary theories of social perception explain them within a predictive processing framework. They argue that perception of others' actions is shaped by prior assumptions about their goals and intentions and the behaviours that these mental states predict. This thesis aimed to resolve whether people make such predictions, whether they are represented perceptually, and on which information they rely.

Ten experiments utilized a variant of the classical Representational Momentum paradigm. They presented participants with the initial stages of a goal-directed action and asked them to make spatial judgments of its last seen position prior to sudden offset. As expected, the results revealed the top-down expectations that guide action perception. The findings revealed (1) that social predictions follow the principle of efficient action, biasing perception towards efficient action expectations, such that hands seen to reach straight towards an obstacle were perceptually lifted over it. These predictions were (2) derived spontaneously, were (3) perceptually represented, and emerged (4) from attributions of intentionality to the observed

actor, even (5) when the action was already underway, based on the match between action kinematics and available target objects.

The current findings provide direct evidence for predictive models of social perception. They confirm that the perceptual representation of others' actions is biased by the intentions we assign to them and our predictions of how these intentions will be fulfilled, therefore providing new avenues to understand how action expectations can shape our understanding of other people's actions.

Table of Contents

List of Figures

1 Chapter One - Understanding the actions of others

1.1 The importance of action understanding for social interaction

Understanding the actions of others is crucial for all social interactions. Through action observation, humans can derive the goals, attitudes and beliefs that drive the actions of others to gain insight into the meaning of their action, and plan a suitable response when required (Bach, Bayliss, & Tipper, 2011; Hamilton, 2009; Sebanz & Knoblich, 2009). For example, when approached in the town centre with a clipboard we know to immediately change direction, when our child hears the jingle of the ice-cream van we are already prepared for the chase into the road, and when we see our friend walking straight-towards the glass door we know he has not seen it.

Deciphering the mental states of others is no mean feat, as they are largely hidden from view. A window into the minds of others must therefore rely upon not only our prior knowledge of the person, the action, or the context (Bach & Schenke, 2017; Press, Heyes, & Kilner, 2011), but also our sensitivity to cues that may indicate the potential goals and beliefs of others. Such cues include where they might be looking (Frischen, Bayliss, & Tipper, 2007; Hudson, Liu, & Jellema, 2009), which emotional expression they have (Hudson & Jellema, 2011), what we already know about their particular preferences (Schenke, Wyer, & Bach, 2016), or even what potential goal objects are available and our knowledge about these objects (Bach, Knoblich, Gunter, Friederici, & Prinz, 2005a; Bach, Nicholson, & Hudson, 2014). In the examples above, we know the intention of the clipboard-holder through our previous encounters, we realise our child has heard the ice-cream van when her gaze suddenly follows it, and we foresee our friends impending misfortune when he does

not slow down upon approach of the glass door (and we know that glass doors cannot be penetrated - without pain).

Utilising these, often subtle, cues is therefore essential for not only inferring the contents of others' minds, but may also provide a stepping stone for developing more sophisticated mentalizing abilities to reason about others' mental states and coordinate social interactions (Hamilton, 2009; Wellman & Brandone, 2009). Understanding the goals, beliefs and attitudes of others can reveal how they may differ from our own, and inform expectations about likely forthcoming actions, guiding our perception of their action and allowing for the preparation of a response to potential action consequences, either cooperatively through joint-action or competitively which would have been crucial for our ancestor's survival (Hamilton & Grafton, 2007; Sartori, Becchio, & Castiello, 2011; Sebanz & Knoblich, 2009).

Action understanding therefore lies at the heart of social interaction. Yet, despite this critical importance, the mechanisms underlying such abilities are still largely unknown or, at the least, controversially debated. It remains unclear what information people use to generate assumptions about other people's goals and intentions, and how this information is translated into an expectation of their upcoming action. Furthermore, little is known about how these expectations guide our perception of their action and contribute to the meaning that we assign to their behaviour. If social perception is influenced by our preconceptions of others, then this leaves open the door for misinterpretation to distort action perception. Uncovering these processes will not only contribute towards our understanding of how humans so effortlessly find meaning in other people's behaviour, but also where these mechanisms breakdown

and are responsible for the marked social deficits in autism spectrum conditions and schizophrenia (Koster-Hale & Saxe, 2013; Pellicano & Burr, 2012).

1.2 Mechanisms underlying action understanding: a bottom-up approach

People's ability to decode the actions of others and assign to them meaning and purpose has conventionally been conceptualised as a bottom-up process (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese & Sinigaglia, 2010; Iacoboni, 2009; Rizzolatti & Craighero, 2004). This process is assumed to rely on dedicated neural processes that "directly match" the incoming visual information from action observation to prior knowledge we already hold about these actions (Gazzola & Keysers, 2008; Rizzolatti & Craighero, 2004). They are assumed to not only rely on learned semantic associations from previous experience that link overt behaviours to their hidden meaning (e.g., that a smile indicates happiness; Brass & Heyes, 2005; Keysers & Perrett, 2004), but to take on a more active role. Simulation theories suggest that we internally embody the actions of others, playing-out the action in our mind as if it were our own (Gallese & Goldman, 1998; Jeannerod, 2001; Wilson & Knoblich, 2005). A dominant view is that these simulations are motoric in nature, simulating the kinematic features of the action in our own motor systems to derive the mental state causes and interoceptive consequences of these actions (Rizzolatti & Sinigaglia, 2010). These simulations are thought to activate the low-level motor areas that correspond to the observed actions, which can then propagate up the observers own motor hierarchy in a bottom-up manner, to reveal the likely goals that would have generated them. For example, when watching someone reach for a hot oven tray, the action is traced onto our own motor system to help us

understand the goals (to move the tray) and beliefs (that the tray is cool enough) behind the action, as well as their potential (sensory) consequences (pain). By simulating the actions of others in our own minds, we can therefore gain insight into the minds of others, drawing on our own motor experiences.

This bottom-up motoric view of action understanding is supported by evidence that action observation facilitates simultaneous execution of the same action. In a seminal study, Brass and colleagues found faster reaction times for executing finger movements when simultaneously observing an actor performing the same finger movements, and slower reaction times (and more errors) when observing different finger movements (Brass, Bekkering, & Prinz, 2001). Since then, various other studies have confirmed this finding for different hand action features such as hand configurations (Stürmer, Aschersleben, & Prinz, 2000), reach directions (Bach, Bayliss, & Tipper, 2010) and even different body parts (Bach, Peatfield, & Tipper, 2007; Gillmeister, Catmur, Liepelt, Brass, & Heyes, 2008; Wiggett, Hudson, Tipper, & Downing, 2011). Observing an action therefore activates similar processes required for the execution of that action, facilitating imitation of congruent actions and interfering with performing incongruent actions (Bertenthal, Longo, & Kosobud, 2006; Press, Bird, Walsh, & Heyes, 2008). Similar effects resulting from measuring motor cortex activations via TMS show that this motor simulation follows a similar time course as the observed action (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995).

Perhaps more compelling evidence for this motoric matching hypothesis comes from the discovery of a group of neurons in monkeys that fire both when the monkey is executing an action and also when merely observing this same action. These “mirror neurons” were recorded in the ventral premotor cortex (area F5; Di Pellegrino et al.,

1992), and later in the inferior parietal lobule (Gallese, Fadiga, Fogassi, & Rizzolatti, 2002), of the macaque monkey brain using single cell recordings. Importantly, mirror neuron activation was only found for goal-directed actions, such as reaching and grasping objects, than for pantomimed (without the object) or intransitive (e.g. tongue protrusion) actions (Di Pellegrino et al., 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996a). This matching between observation and execution is therefore specific to actions from which action meaning can be deduced. These findings revealed, for the first time, a neural mechanism that may directly translate the visual information about a seen action into the motor command that would produce it. The observer likely knows the potential sources of this command, and so allows them to understand what the actor is doing, and, perhaps, why they are doing it.

Evidence that mirror neuron activity represents action understanding comes from studies that show firing for actions that grasp an object in full view, as well as for actions that grasp an object which is hidden behind an occluder (Umilta et al., 2001). Importantly, mirror neuron firing for hidden grasps was only evident when the monkey had previous exposure to the hidden object. This firing was almost eliminated when the monkey did not know what was behind the occluder. Therefore, the observed action triggered the equivalent action in the observer's own motor repertoire only when the observer could make use of prior knowledge to understand the action. This finding is supported by further studies, which found mirror neuron firing for previously known and distinctive action sounds (e.g. breaking nuts or tearing paper), even in the absence of visual information (Keysers et al., 2003; Kohler et al., 2002). Therefore, as long as the action can be recognised, by prior knowledge of object placement or by knowledge of action sounds for example, mirror neurons can generate the motor plan required for their execution.

Following this ground-breaking discovery, attempts have been made to identify equivalent systems in the human brain, albeit using more indirect neuromodulation methods such as transcranial magnetic stimulation (TMS; Aziz-Zadeh, Maeda, Zaidel, Mazziotta, & Iacoboni, 2002; Fadiga et al., 1995; Gangitano, Mottaghy, & Pascual- Leone, 2004) and neuroimaging methods such as functional magnetic resonance imaging (fMRI; Aziz-Zadeh, Koski, Zaidel, Mazziotta, & Iacoboni, 2006; Jonas et al., 2007; Shmuelof & Zohary, 2006). Indeed, it has been shown that mirror systems do exist in the human brain, in areas homologous to that of the monkey mirror system. Activation in ventral premotor cortex (PMv), inferior frontal gyrus (IFG) and inferior parietal lobule (IPL) has been found for the observation and execution of comparable actions (Rizzolatti & Craighero, 2004; Rizzolatti & Fabbri-Destro, 2008). Furthermore, corresponding activation for action observation and execution has been identified at the single voxel level in humans (Mukamel, Ekstrom, Kaplan, Iacoboni, 2010).

Further evidence for a mirror system in humans comes from studies comparing mirror activity in response to actions that an individual was an expert at performing, with individuals who had less experience (Cross, Hamilton, & Grafton, 2006; D'Ausilio, Altenmüller, Olivetti Belardinelli, & Lotze, 2006). Since the mirror system is assumed to rely on prior knowledge to simulate the corresponding motor activation, it can be predicted that experts would show the greatest mirror activation. Indeed, brain imaging studies showed that expert piano players and expert ballet dancers produced greater mirror activation for piano key presses and ballet dance moves, respectively, than non-experts (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2004; Haslinger et al., 2005). The prior knowledge and experience of performing these actions therefore elicited a stronger response to their observation

compared to those with less experience, activating the motor processes involved in their execution that the experts would be all too familiar with.

Unlike what is known so far about the monkey mirror system, humans also appear to mirror non-goal-directed and intransitive actions, such as gestures or meaningless arm movements (Grèzes, 1998; Grèzes, Costes, & Decety, 1999; Press et al., 2008). Rather than being evidence against the role of the human mirror system in action understanding, it instead proposes an important difference between humans and primates when watching the actions of others (Press et al., 2008). The human ability to attribute meaning to seemingly goal-less arm movements essentially reveals the unique ability to identify symbolic communication, as well as allowing for learning through imitation. In humans, non-goal directed hand actions are used during gesturing and sign language, where the goal is not to act upon an object but to communicate. By mirroring these actions, humans can access their symbolic meaning and facilitate the understanding of this action (Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004). The mirror system is also argued to support action understanding in language not only through the use of gestures or sign, but also by using spoken word instead of observed action as the sensory input. Prior knowledge of these action words can then map the action information onto the motor programs that correspond to the action, facilitating the understanding of the action words (de Lafuente & Romo, 2004; Hauk, Johnsrude, & Pulvermüller, 2004).

In summary, the literature reviewed here provides evidence for a mirror system in humans that plays a central role in simulating the actions of others within the observers own motor system. This account explains the neural mechanisms that serve the transition of action information from its sensory source into the motor

components responsible for its execution. This ability to embody the actions that we observe, hear, or even talk about, is thought to enable us to understand the meaning behind the actions, from the inside. Simulating the kinematic properties of these actions, as if we were performing them ourselves, allows us to draw from our own personal experiences to relive not only the performance of the action, but to also reactivate their likely causes and/or consequence.

While the literature provides convincing evidence for action simulation and the mechanisms responsible for translating the observation into motor representation, they do not provide direct evidence that these mirror activations support the actual understanding of the action. Recent findings have come to light that cannot be fully explained by motor activation alone, challenging the purely bottom-up account and suggesting an involvement of top-down processing. It has been suggested so far that the mirror system has some reliance upon prior knowledge and experience when assigning meaning to observed actions. The following sections will discuss the pitfalls in the bottom-up account of action understanding and suggest an alternative account of how we so effortlessly navigate the social world.

1.3 Problems with the mirror system account of action understanding

The ability to understand the actions of others comes naturally to humans. Despite the largely hidden drivers of these actions, we are able to effortlessly deduce the likely goals and mental states from which these actions were derived (Hamilton & Grafton, 2007; Sebanz & Knoblich, 2009). The bottom-up account of action understanding, as outlined in the previous section, describes a mechanism in which observed actions are directly translated into the motor components responsible for

their execution (Gazzola & Keysers, 2008; Rizzolatti & Craighero, 2004). Activation of these low-level motor features can propagate to higher motor levels in the observer's own motor system, revealing the goals and beliefs that would have driven those same actions in the observer. The meaning of the action is therefore understood by relying on the observers own motor experience of the observed action (Rizzolatti & Sinigaglia, 2010).

This explanation of action understanding has an intuitive appeal, especially considering the supporting evidence from mirror-neurons in monkeys and homologous mirror systems in humans as described above (Di Pellegrino et al., 1992; Gallese et al., 2002). However, this explanation fails to consider that there is no one-to-one mapping between actions and their goals. The same action can achieve a number of different goals and the same goal can be achieved using numerous different actions (Bach et al., 2005a, 2014; Jacob & Jeannerod, 2005). For example, a reach towards a cup could have the goal to drink from its contents or to move it for cleaning. Indeed, Iacoboni et al. (2005) presented these very actions to human participants, with and without the two different contexts. Increased activation was found in areas of the human mirror system when the action had context (depicted by other objects in the background), compared to when the action was presented in isolation. The authors claimed that this is evidence for a motoric basis of action understanding since the mirror system was more strongly activated when the meaning of the action could be deduced. However, the action kinematics were identical in both conditions, and only the context was different. It is therefore more likely that this difference reflects action understanding from outside the motor system, that encodes the context, and not from the observation of the action alone. Consequently, it seems quite impossible that a single action can be understood in a

purely bottom-up motoric matching way, without relying on additional sources of information from higher regions that need not be motor related (Hickok, 2009; Jacob & Jeannerod, 2005).

Other evidence also suggests that findings previously thought to support a motor simulation account of action understanding may actually serve as evidence against it. As discussed in the previous section, Umilta and colleagues (2001) suggest that the mirror neuron activation for actions with occluded targets aids action understanding, or at the very least, action recognition, since this only occurs when the observing monkey is aware of the occluded object. However, the monkey has the same visual access to motor information both when the occluded target is known and when it is unknown. Therefore, prior knowledge of the occluded target is what contributes to action understanding, rather than the observed action kinematics. This understanding may, in turn, trigger the differences in mirror activation, suggesting that mirror neurons may *reflect* the understanding of an action that was already achieved by other means, rather than being *causally responsible* for it.

Alternative views of the role of the mirror system may also be required to explain recent findings from lesion studies that contradict traditional bottom-up theories of action understanding. A motor simulation account would predict parallel disruption to action execution and observed action understanding with damage to mirror regions. Nevertheless, while some lesion studies do reveal such corresponding impairments, others do not, creating inconsistencies in the literature. Buxbaum and colleagues (2005) studied 44 patients with left-hemisphere damage and found strong correlations between performance on action recognition and execution tasks. These findings were replicated by Negri et al. (2007) who found similar correlations with

equally large samples of comparable patients, providing evidence consistent with the direct-matching hypothesis. However, when patients were analysed at an individual level, Negri et al. (2007) uncovered dissociations between the ability to recognise actions and the ability to execute them. Moreover, in some studies, action comprehension deficits have been associated with damage specifically to the IPL (Weiss, Rahbari, Hesse, & Fink, 2008), while others demonstrate this impairment in patients with IFG lesions (Pazzaglia, Pizzamiglio, Pes, & Aglioti, 2008). These contradictory findings therefore make it difficult to ascertain the specific role of the human mirror system and how crucial the motor system is for action understanding.

Moreover, a reliance upon one's own motor system for action understanding implies that one must possess the motor knowledge and experience required to perform the observed action in order to understand it. This would therefore be a problem for infants who have not yet mastered the ability to execute certain actions that they see others perform, for disabled individuals who may have similar action execution deficits, and as well, for novices observing expert performers. When watching the skilled footwork of a football player as he tackles another player, we understand that his intention is to gain possession of the ball, despite our inability to perform such actions ourselves. Moreover, we often have no difficulty in understanding the actions of animals that are biologically impossible for humans to perform, such as birds flying to their nests. Such examples provide further challenges for the direct-matching hypothesis, further emphasising a role for systems outside of the motor system (Hickok, 2009).

Together, these findings suggest that the role of the human mirror system in action understanding may not be so clear-cut. While evidence for this bottom-up pathway

undeniably exists (di Pellegrino et al., 1992; Gallese & Sinigaglia, 2010; Iacoboni, 2009; Rizzolatti & Craighero, 2004), the inconsistencies and irregularities of the research findings discussed here are enough to cast doubt upon this explanation of action understanding. The field of social perception is therefore forced to seek an alternative theory of how humans are able to extract the hidden goals and mental states that guide the action of others, facilitating all social interaction. One potential avenue to explore, as hinted earlier in this section, may point to a role-reversal between mirror system activity and action comprehension. Such an idea would turn the bottom-up approach to action understanding on its head. Rather than motor simulation providing the information needed to comprehend the actions of others, this might instead be a consequence of action understanding, signifying when the action has already been understood, and is now tested against the perceptual input (Csibra, 2008; Mahon & Caramazza, 2008). Evidence for a top-down approach to action understanding is discussed in the following section.

1.4 An alternative account of action understanding: Top-down prediction.

Some of the findings discussed in the previous section move away from a purely bottom-up approach to action understanding and allude to a top-down mechanism in which mirror neuron activation is secondary to action understanding, as suggested by several authors. In such views, mirror activation reflects a *consequence* of action comprehension – that is perhaps acquired in other regions - rather than its *cause* (Jacob & Jeannerod, 2005; Mahon & Caramazza, 2008). Such an approach would use alternative mechanisms, rather than purely motor-matching processes, to facilitate the recognition of others' actions and their goals. Once realised, motor

circuits could be engaged as a consequence of this understanding, perhaps to predict which actions might follow from the inferred goals, or to verify that the action was understood *correctly* (Bach et al., 2014; Csibra, 2008; Kilner, Friston, & Frith, 2007ab).

Indeed, recent evidence suggests that mirror activation responds to expected actions, rather than observed actions, indicating a predictive process as an underlying mechanism for action comprehension. Fogassi et al. (2005) found activation in different mirror neurons when a monkey watched an experimenter initially grasp an object with the goal to eat it, compared to grasping the object with the goal to move it. Importantly, these mirror neuron activations were comparable to those elicited when the monkey executed the respective actions. This suggests that, even though the observed actions towards the object were the same, the goal of the action could be inferred from available goal cues (e.g. the presence of a container), and so the mirror neuron activation may well reflect the anticipated action that would have continued beyond the initial reach to grasp. Similar findings were reported by Maranesi and colleagues (2014), whereby the onset of mirror neuron activation for predictable actions preceded action observation by 340ms.

These findings are mirrored by recent studies in humans. In a behavioral study, Liepelt, Von Cramon, and Brass, (2008) found that motor activation during action observation reflected the implied goal of an action, rather than the observed action itself. When participants observed an actor attempting to raise their finger, despite being restrained by a finger clamp, motor responses of the participants reflected finger lifting, even though this was not observed. Further evidence from adult and from infant studies using EEG to investigate the time course of action mirroring found

motor activation prior to observing the action when the action was predictable, compared to when goal information was unavailable (Bozzacchi, Spinelli, Pitzalis, Giusti, & Di Russo, 2015; Fontana et al., 2012; Kilner, Vargas, Duval, Blakemore, & Sirigu, 2004; Southgate, Johnson, Osborne, & Csibra, 2009).

In these examples, the goal of the action can already be assumed from prior knowledge or context, which can inform a prediction of the action required to achieve this goal. It is therefore more likely to be the *expected* action that is simulated, rather than the *observed* action. Such a mechanism indicates a major role for other brain systems in action understanding, moving away from a reliance upon motor activations, which may instead reflect the consequence of action understanding (Jacob & Jeannerod, 2005; Mahon & Caramazza, 2008).

Indeed, such claims were made by Csibra (2008) who challenged the bottom-up motor-matching account, suggesting that this “simulation” of observed actions actually reflects “emulation” of expected actions. This action reconstruction account argues, first, that potential action goals and intentions are derived outside of the motor system by relying on prior knowledge through associative learning and teleological reasoning. Instead of mirroring observed actions to reveal their goals, goal estimations are made, in a Bayesian manner, based on the statistical probability that an action outcome is generated by a specific action and the probability that the action is generated to fulfil a specific goal. These probabilities are not only informed by previous experience with the action as well as potential target objects, but also by reasoning that humans attempt to achieve their goals in the most energy efficient way as possible, accounting for potential contextual constraints (e.g., location of goal objects, obstacles etc.). Csibra argues that these goal assumptions then generate

predictions about the actions that would be required to achieve these goals if they are correct. From this view, motor activation does not serve to reveal action goals, but instead reveals a search for an action that would achieve the inferred goal and compares this with the observed action as it unfolds. A close enough match would confirm the inferred goal and a mismatch would generate a revision to the emulated action. This account therefore provides a predictive mechanism for social perception, with motor activation reflecting emulations of expected actions, and goal understanding emerging from systems outside of the motor domain.

Csibra's action reconstruction account is very much in line with the predictive motor activation literature discussed above. It suggests that, in Liepelt et al.'s (2008) finger clamp study for example, the goal of the action (to lift finger) engaged the observer's motor system to emulate the action that would achieve this goal (raising the finger up), rather than simulating the observed action (finger failing to raise up). This account can also explain findings from similar studies in infants. Motor activation was only found for predictable actions with an obvious goal (a grasping hand reaching behind an occluder), and not when the same action could not be well understood (a mimed grasp; Southgate, Johnson, Karoui, & Csibra, 2010). Moreover, Bach and colleagues (2010) reported action mirroring effects only when the observed grip type matched the size of a target object.

This top-down approach was developed further by Kilner, Friston, and Frith, (2007ab), by drawing parallels between the predictive nature of social perception and perception in general, which is understood within a predictive coding framework (Clark, 2013; Friston, 2010). In this view, actions are represented across many interconnected levels of an action hierarchy. The highest level describes the

intention of an action (e.g. to eat), the next level describes the series of short-term action goals that are required to fulfil the intention (e.g. reaching for an apple and bringing it to the mouth), and the lower levels describe the action kinematics required to achieve these short-term goals (e.g. the path of the reach and the grip formation; (Hamilton & Grafton, 2007). Each level of the hierarchy forms a prediction about the representation in the level below. This prediction is then compared to the representation and any difference is communicated back up to the level above in the form of a prediction error. The representation in the higher level is then adjusted to better explain the lower-level representation, generating new predictions about these lower-level representations, and this process continues until the prediction error is minimised at all levels of the hierarchy.

Generally, therefore, in both Kilner's and Csibra's views, social perception, like perception in general, can be understood as a process of Bayesian hypothesis testing and revision: knowledge of the world and its statistical regularities is translated into probabilistic predictions about the perceptual input that would be received if assumptions about the world are correct (Bach & Schenke, 2017; Csibra, 2008; Kilner et al., 2007ab). As incoming sensory evidence is received, hypotheses are constantly updated until the "best guess" that explains the evidence is arrived at, and perception reflects this current best estimate of the sensory input. Social perception is therefore seen as an attempt to bring, via predictions and prediction errors, higher- and lower-levels of such action hierarchies into alignment.

Kilner and Csibra's proposal align models of social perception with more general predictive processing accounts that have been used to explain perception across different domains, with evidence from visual, auditory and motion perception to name

but a few (as will be discussed in more detail in the following section). These models explain not only how we represent the world around us, but also how these representations can be biased towards what we believe to be true. The following sections will discuss evidence for the predictive processing account of perception and how social perception in particular can also be understood as a top-down process.

1.5 Evidence for predictive processing in non-social perception

Predictive processing models of perception (see above) argue that what we perceive results from a series of hypothesis testing and revision, predicting the incoming sensory information from their inferred cause and updating the inferred cause as the sensory information is received, until the most likely cause of the input is generated. These models explain perceptual processing in sensory domains across the brain and the resulting distortions to perception that are often evidenced by perceptual illusions.

In vision, these models explain how we perceive the “true” colour of a surface, which is not only dependent on the colour information provided by the senses but also on our expectations of the light source and surrounding illumination (Bloj, Kersten, & Hurlbert, 1999). A popular example is the “blue dress illusion” (see Chetverikov & Ivanchei, 2016). In a photograph of a dress, some people report seeing a blue and black dress while others report the dress to be white and gold. Predictive processing accounts explained this illusion in terms of differing beliefs about the surrounding illumination that generated predictions about the colour information that would be received by the senses. Those people that assumed the light source to be from the

front, such as the cool light from a camera flash, perceived the dress to be blue and black, and those that assumed the light source was from behind, such as warm sunlight from a window, perceived the dress to be white and gold.

Predictive coding similarly explains the mechanism underlying bistable figures that seem to spontaneously flip as the brain explores differing hypotheses about the sensory input (Hohwy, Roepstorff, & Friston, 2008). When two different images are presented to each eye, the resulting percept is not a blend of the two images, as one might expect. Instead, each image is perceived separately, one-at-a-time, and the constant flipping between experiences of the two images may reflect the testing of differing hypothesis about what is being presented, and their subsequent partial confirmation from the input of one eye.

Predictive processing models do not only explain perception within one sensory modality, but also the well-documented crosstalk between them. In audiovisual speech perception, visual information from the shaping of the mouth during speech is used to predict the speech sounds that are heard. The McGurk effect (McGurk and MacDonald, 1976) demonstrates that when these two sensory inputs are incongruent, i.e. seeing the mouth movements that predict a particular speech sound but hearing a different sound, speech perception is distorted as the brain tries out a new hypothesis that accounts for both sensory inputs. Moreover, the magnitude of the illusion is increased the when illusory speech sound is embedded into a semantically congruent sentence, compared to an incongruent sentence (Windmann, 2004), reinforcing the influence of expectation.

Important for social perception, top-down influences on perception are also evident in motion perception. Apparent motion describes the phenomenon that illusory

motion is perceived between two visual stimuli that are flashed in quick succession (Kolers, 2013; Wertheimer, 1912). The two stimuli are typically not seen as alternating but instead perceived as one object moving between two different locations. This apparent motion effect reflects the prior hypothesis of a single moving stimulus generating a predicted motion path between the two locations. The perceptual system fills in the predicted but missing steps in-between as the input, albeit partially, confirms this hypothesis (Muckli, Kohler, Kriegeskorte, & Singer, 2005; Sterzer, Haynes, & Rees, 2006). Similar processes explain why perceptual blurring is minimised during motion perception (Bex, Edgar, & Smith, 1995; Hammett, 1997), as top-down inferences sharpen the degraded input.

Predictions during motion perception not only compensate for degraded or missing input during the motion, but also account for future motion, predicting where moving stimuli will continue and perceptually inserting the predicted next step. Freyd and Finke (1984) found that people generally over-estimate the disappearance point of a moving stimulus in the direction of the motion trajectory, an effect known as Representational Momentum (Hubbard, 2005). When asked to compare a probe stimulus to the last seen image before it disappeared, participants were more likely to misperceive probes positioned ahead of the final image (forward probes) as being in the same position, compared to backwards probes. This phenomenon is taken to reflect that prior movement information is used to make constant predictions about where the moving target will continue in following steps. When the input is suddenly removed, a reliance upon these top-down predictions creates a match between the forward probe and the predicted position, whereas backward probes more readily detected and signal a prediction error.

Further studies have then shown that these forward biases do not only capture the low-level predictions of future kinematic movement but also incorporate higher-level expectations from prior knowledge (Hubbard, 2005; Hudson, Bach, & Nicholson, 2018; Hudson & Jellema, 2011; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016; Jordan & Hunsinger, 2008), in line with predictive processing models of perception. For example, larger forward biases were found for vertical movement from top to bottom, compared to bottom to top. This effect, known as representational gravity, therefore provides evidence that the visual system not only uses bottom-up movement information to generate predictions of future positions, but also incorporates high-level knowledge about physics (i.e. that objects fall down but not up, Hubbard, 1997, 2005; Hubbard & Bharucha, 1988). Furthermore, horizontal movement is less future-biased when the target moves along a rough surface compared to a smooth surface (representational friction; Hubbard, 1995, 2005). Together, these Representational Momentum examples therefore provide evidence of a predictive mechanism in motion perception that not only utilises low-level information from the current target motion, but also makes use of high-level information from prior knowledge about how this motion is differentially affected by the environmental context. Moreover, they provide further evidence for a hierarchical structure of predictive processing, with reciprocal communication between high-level and low-level predictions to inform the resulting perceptual representation.

Predictive processing models of perception have provided a robust explanation of the experiences we encounter every day, not only in the visual domain, but across all sensory modalities (Clark, 2013; Csibra, 2008; Friston, 2010; Kilner et al., 2007ab). By understanding perception as the brain's "best guess" of what is being experienced, informed by prior knowledge and context, we can better understand

how we are able to so easily deduce the sensory input we receive despite the ambiguity and noise that comes with it, such as during motion perception or during conflicting sensory information, or occlusion in the visual field (Bex et al., 1995; Hammett, 1997; Kok, Brouwer, van Gerven, & de Lange, 2013; Lages, Boyle, & Jenkins, 2017; Muckli et al., 2005; Sterzer et al., 2006). Moreover, this allows us to describe how our eyes play tricks on us, and how expectations can lead our understanding of sensory information astray, as evidenced by various visual illusions (Bloj et al., 1999; Chetverikov & Ivanchei, 2016; Hohwy et al., 2008). Since predictive processing models have provided such robust explanations of perception spanning several different domains, it would not be ill-considered to assume that similar processing indeed underlie social perception.

1.6 Evidence for predictive processing in social perception

The above reviewed findings provide evidence that perception in general is generated via top-down predictive processing (Clark, 2013; Friston & Kiebel, 2009a). Given the generalisability of these predictive processing theories of perception across multiple sensory domains, it would be logical to assume that social perception, too, relies upon comparable mechanisms. This would suggest that, as for perception in general, social perception is guided by top-down predictions about others' actions and reflects our best guess about what has been observed (Bach et al., 2014; Bach & Schenke, 2017; Csibra, 2008; Kilner et al., 2007ab). Action understanding would then similarly not only depend upon what is observed, but also upon what was expected. While such predictive models have only recently been developed, there is already some early evidence for them.

Early evidence for social prediction emerged from studies measuring predictive eye movements during action observation. When watching an actor performing an action, such as stacking blocks, the observer's pattern of eye movements mirrored the pattern of eye movements elicited when they performed the action themselves (Flanagan & Johansson, 2003). Importantly, these eye movements were not reactive to the observed action but were predictive. They were directed towards the goal object in anticipation of the actor's reach towards it, suggesting that eye movements are guided by top-down knowledge about the actor's intentions rather than driven by a response to the observed action (Eshuis, Coventry, & Vulchanova, 2009a).

Evidence for predictive social perception has also been found in studies of apparent motion. Originally designed to measure top-down prediction in motion perception (see previous section), these studies replaced object stimuli with biological stimuli to demonstrate how similar predictive mechanisms are at play (Chatterjee, Freyd, & Shiffrar, 1996; Shiffrar & Freyd, 1990,1993). As with object stimuli, flashing an image of an arm or leg at two different locations in rapid succession created the perception of motion between these two positions. However, motion perception for human action was specifically modulated by human kinematic capabilities and action goals, with adjustments to perceived motion paths when the shortest distance between the two locations was either obstructed or was not humanly possible to perform. This demonstrated that social perception relies upon similar mechanisms as motion perception in general, and is additionally influenced by prior knowledge about the specific stimulus' properties of human bodies and how they can move and not move.

These studies not only provide support for a predictive processing account of social perception, but also suggest that social perception may be similarly susceptible to

confirmation bias, favouring interpretations of the input that provide a close-enough match to prior assumptions. This creates the opportunity for distortions in perception, akin to the visual illusions described earlier, especially when the incoming sensory information is ambiguous (Kok et al., 2013; Lages et al., 2017). For example, a quick glimpse of a hand forming a large grip might appear to be closer to a large target object (such as an apple), compared to a nearby smaller object (such as a strawberry), than it really was, in line with the prediction that the actor wants the apple (i.e, affordance matching, Bach et al., 2014). In this example, the quick glimpse of the action may only generate limited and ambiguous visual information, forming an interpretation in light of the expectation that the apple will be grasped, therefore imbuing the observation with meaning. Perception, in such models, would therefore not be veridical, but potentially influenced by our prior expectations, especially when sensory information is limited or degraded (Kok et al., 2013; Lages et al., 2017).

Evidence for predictive biases in social perception is reported in motion perception studies that have been adapted to measure action prediction. Representational Momentum studies, as described above, reveal predictive biases towards the future position of a moving target. When tested with biological motion such as arm reaches and head rotations, perception of these actions were similarly biased towards future positions along the motion trajectory (Hudson, Burnett, & Jellema, 2012; Hudson et al., 2009; Hudson & Jellema, 2011). As with the apparent motion studies above, these perceptual biases were modulated by prior knowledge about the actions' goal. Rotating heads were perceived to have rotated even further when the eyes were gazing towards the same direction, than when they gazed towards the opposite direction. Here, the direction of the eyes were used as a goal cue to inform

predictions about how the rotation of the head would continue, creating a bias to the perception representation of the action.

Similarly, arm reaches were reported to have reached even closer to target objects when the actor had announced, “I’ll take it!”, than when they announced, “I’ll leave it” and vice versa when the arm withdrew away from the target object (Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, et al., 2016). The degree to which perceptual judgements of arm movements were distorted in the direction of motion was not only modulated by the prior knowledge of the actors intention, but also by how predictive this intention was of the action (Hudson et al., 2018). The perceptual bias towards the actor’s intention was larger when the actor was more likely to perform the corresponding action, than when they were more likely to perform the opposite action (e.g. saying “I’ll take it” and then withdrawing from the target object).

Together, these studies provide tantalizing evidence of the predictive nature of social perception, in line with predictive processing models of perception in general. These perceptual biases for human action do not only follow similar mechanisms as motion perception, but are specifically informed by prior information about how humans typically behave (Chatterjee et al., 1996; Shiffrar & Freyd, 1990, 1993). These studies demonstrate that prior information is gathered from biological sources, such as eye gaze direction and speech, as well as from the environment, such as available target objects and obstructions, to cue the potential goals of the action (Bach et al., 2014; Bach & Schenke, 2017; Hudson et al., 2018; Hudson & Jellema, 2011; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016). In line with Bayesian integration and predictive coding models of perception,

this prior knowledge is translated into a weighted prediction about how the action will continue if our goal assumptions are correct. The resulting perceptual biases in line with expectations suggest that high-level goal information cascades down to lower perceptual levels, guiding the sensory input and therefore biasing the perceptual experience of other's actions. As in perception in general, social perception is therefore not veridical but instead reflects a best guess of the actor's goal and represents a prediction of how this goal will be achieved.

While these studies provide compelling evidence for comparable mechanisms between social perception and non-social perception, offering a revolutionary perspective on how we understand the actions of others, the findings are limited and so many questions about the underlying mechanisms of action understanding remain unanswered. It is still unclear how these predictions of others' upcoming behaviour are generated, on what cues they rely, how these cues are weighted and integrated to inform these predictions, and how these predictions alter the perceptual representation of others' action, for example.

1.7 Thesis Overview

This thesis aims to resolve which information is used to inform action predictions and what impact these predictions have on social perception, therefore providing new avenues to understand how action expectations can shape our understanding of other people's actions. The experiments presented in this thesis utilise a modified version of the Representational Momentum paradigm (Freyd & Finke, 1984; Hubbard, 2005) to capture the predictions made about biological actions and to measure their resulting effect on action perception. Participants will be presented

with a series of short action clips that contain cues to the actor's intention, allowing for the formation of action predictions prior to or during action observation. The action will suddenly disappear and participants will be required to indicate its final location. Comparing these responses to the real final position allow for the direct measure of perceptual distortion, capturing the subtle influences of action predictions on action perception. This novel paradigm goes one step further from original Representational Momentum studies in action observation that use probe stimuli to measure perceptual biases, by instead employing touch-screen responses to pinpoint more specifically that magnitude of perceptual distortion. A detailed outline of this methodology is presented in Chapter Two.

The experiments in Chapters Three, Four and Five utilise this Representational Momentum paradigm to probe action understanding and to uncover a potential cue that may inform action predictions. As described earlier, Csibra (2008) identified three sources of information, derived from outside the motor system, that indicate the goals and intentions of others actions. As well as relying on associative learning to derive the statistical probability that an action was executed to achieve a particular goal, and the probability that such a goal could be achieved with the particular action, a main source contribution comes from the understanding that humans generally act efficiently to reach their goal. This assumption of efficient action proposes that humans take the most direct path to achieve their goals, minimising energy expenditure and time, and only exerting additional energy when the environment requires this – such as when obstacles in the way need to be overcome (Csibra & Gergely, 2013; Dennett, 1987; Gergely & Csibra, 2003).

Previous research suggests the assumption of action rationality is a basic building block of human social perception skills. It has been previously shown that even infants and some primates make assumptions about action efficiency, showing surprise when a rational actor does not act efficiently, by not attempting to avoid obstacles or by taking an unnecessarily long route to their goal, for example (Gergely & Csibra, 2003). Once established, this simple heuristic of efficient action may contribute to more sophisticated theory of mind abilities (Wellman & Brandone, 2009). For example, observing an actor reaching inefficiently towards an obstacle may signal that the actor has a different belief to our own (i.e. that there is no obstacle). Indeed, viewing such actions has been shown to activate perceptual and mentalizing processing in the brain (Desmet & Brass, 2015; Marsh, Mullett, Ropar, & Hamilton, 2014).

The experiments in Chapter Three harness the assumption of action efficiency within a Representational Momentum paradigm to test (1) whether action efficiency is used to inform predictions, (2) whether these predictions specifically concern the expected action kinematics in a perceptual format, (3) how these predictions bias the actions' perceptual representation, and (4) whether these perceptual biases are elicited spontaneously. They investigate how perception of efficient action is distorted by adding or removing obstructing objects into the path of a reaching arm, rendering the action inefficient. If specific action predictions are made based on action efficiency, and in a perceptual format, we would expect the perceptual representation of inefficient actions to be distorted towards efficient action trajectories. For example, perception of an arm reaching straight towards an obstruction should be shifted upwards, in line with the prediction that the arm will lift to avoid it. Similarly, perception of an arm reaching over an empty space should be shifted downwards, in

line with prediction that the hand will take a more efficient, straight path. Importantly, these perceptual shifts should be elicited when passively viewing the actions, but should increase with increasingly explicit instructions to predict or take account of the environmental constraints on the action.

The experiments in Chapter Four further investigate the perceptual nature of action predictions by testing whether the effects found in Chapter Three do indeed reflect changes to the perceptual representation of the observed actions, rather than effects produced by motor or working-memory components. They investigate whether these perceptual biases (1) generalise to probe judgement tasks that do not require working memory or touch screen responses and (2) can be disrupted by dynamic visual noise masks that disrupt lower-level perceptual processing. If substantiated, these findings would provide more support for a perceptual locus of action predictions.

The experiments in Chapter Five investigate how these predictions of action efficiency are generated. They test the specific features of the action stimuli that predictions of action efficiency rely on. As described earlier, predictions of action efficiency rest on the reasoning that humans are intentional agents who generally act in the most energy-efficient way as possible. Removing the cues that signal the intentionality of the agent should therefore reduce the formation of efficiency predictions. In these experiments, intentionality cues were manipulated by replacing the actor's arm with a non-agentive ball that still moved with the same biological profile as the arm. A further experiment removed a further cue to intention by removing this biological profile so that the ball now moved in a straight line at a constant speed. If predictions of action efficiency rely on cues to intention, then the

size of the perceptual effect should be reduced when the arm is replaced by a ball, and reduced even further when the biological profile is removed.

All experiments in previous chapters provide enough information in the initial stimulus frame to make a prediction before action onset (e.g. by presenting a goal object and an obstacle that has to be reached over). The experiments in Chapter six test whether action predictions are elicited online, during the unfolding action. Here, only presenting cues to inform goal assumptions once the action is already underway will reveal whether action predictions are indeed generated in real time, during ongoing action observation. To test this, these experiments make use of the assumption that the available objects in a scene and our knowledge about them can provide cues to others' goals and intentions (see Bach et al., 2014 for a review). Knowledge of an object's function and how they are typically manipulated can inform predictions about how an intentional actor, who wants to interact with the object, would achieve this. For example, a small object affords a small grip. To reach and grasp a small object, an actor would be expected to form a precision grip – with the index finger and the thumb – in order to successfully grasp the small object in the most efficient way. Observing the formation of an actor's grip and identifying the match to potential goal objects can therefore signal the intention of a rational actor. The experiments in this chapter investigate (1) whether action predictions are made when goal information is only available during the ongoing action, (2) based on the match between the unfolding grip-formation and affordance of available objects, (3) how these predictions bias the actions' perceptual representation, and (4) if these biases are elicited automatically. These experiments test whether predictions are made by the unfolding match between grip-formation and object affordance, and whether these predictions rely on explicit goal assumptions. If perception of specific

grip formations are distorted towards compatible compared to incompatible objects, and are present even when goal assumptions are implicit, then we can confirm that humans make on-line adjustments of predicted actions based on the match between hand grip and object goals, distorting the perceptual representation of the action, and that these distortions may not reflect high-level goal assumptions, but emerge from relatively low-level processing of kinematic features within the perceptual system.

2 Chapter Two – Using Representational Momentum to measure action prediction

2.1 Representational Momentum

This thesis aims to resolve the processes that govern predictive social perception. The current work applies predictive processing theories (e.g., Clark, 2013; Friston, 2010) to social perception to investigate how we form expectations about other people's behaviour, upon what information these expectations rely, and how they ultimately shape our perception of other people's actions, contributing to action understanding. Specifically, the experiments in this thesis measure the influence of action predictions on social perception to reveal how our observations of other people's behaviour may not be veridical, but biased by what we expected them to do. They utilise the Representational Momentum paradigm (Freyd & Finke, 1984; Hubbard, 2005), originally designed to measure motion perception in general, to capture the predictions made about biological actions and to measure their resulting effect on action perception. This chapter will briefly review the Representational Momentum paradigm as a measure of predictive displacement, in line with predictive processing theories of perception, and how this method can be applied to social prediction, providing a novel tool to reveal how action expectations influence social perception.

2.2 What is the Representational Momentum effect?

In 1984, Freyd and Finke presented to participants three successive presentations of a rectangle in progressively increasing orientations to imply rotation, followed by a fourth "probe" rectangle. Participants were asked to compare the orientation of the

probe rectangle to the final presentation of the rotating rectangle. They found that probes that were rotated slightly further in the direction of motion than the last seen image were more likely to be incorrectly judged as “same” than probes rotated not as far. The authors named this effect the Representational Momentum effect, which is more commonly defined as the over-estimation of the disappearance point of a moving stimulus in the direction of the motion trajectory (Freyd & Finke, 1984; Hubbard, 2005, 2017; Thornton & Hubbard, 2002). More generally, when asked to compare a probe stimulus to the last seen image before it disappeared, observers are more likely to misperceive probes positioned ahead of the final image in the direction of motion (forward probes) as being in the same position, compared to backwards probes. The Representational Momentum effect therefore represents the anticipation of the future movement of a target and has been observed for targets moving horizontally (Hubbard, 1995; Hubbard & Bharucha, 1988), vertically (Hubbard, 2001; Nagai, Kazai, & Yagi, 2002), inwards on a curved trajectory (Freyd & Jones, 1994; Hubbard, 1996) and even for implied motion in frozen action photographs (Freyd, 1983; Freyd & Pantzer, 1995).

The term “Representational Momentum” was coined from the original hypothesis that the representation of a moving target incorporated the physical principle of momentum, following the same rules of physics as the moving target itself (Finke, Freyd, & Shyi, 1986). For example, the eyes still continue to move for a time in the direction of the target once it has disappeared, in the same way that a moving target cannot immediately stop when resisting force is applied (Hubbard, 2017; Thornton & Hubbard, 2002). Indeed, a number of studies have identified physical variables that influence the magnitude of the representational momentum effect, in line with the theory of an internalisation of physics. For example, target objects moving

downwards, in the direction of gravity, elicit a larger representational momentum effect than objects moving upwards, against the force of gravity (Hubbard, 1995, 1997; Hubbard & Bharucha, 1988). Similar influences of implied physical factors have been shown to affect representational momentum, such as friction acting on the target (Hubbard, 1995b, 1998), the weight of the target (Hubbard, 1997b), acceleration of motion (Finke et al., 1986) and velocity (Freyd & Finke, 1985).

More low-level theories of representational momentum suggest that the effect is elicited by smooth pursuit eye-movements, and represents ocular overshoot after target offset (Kerzel, 2000; Kerzel, Jordan, & Müsseler, 2001). Kerzel (2000) argued that the visual persistence of a target on the retina continues for 50-60ms after its offset and together with the continuing movement of the eyes, creates a distortion in the direction of motion. However, smooth pursuit eye-movements only result from observing smooth motion, and so cannot explain forward displacement for apparent or implied motion (Hubbard, 2017).

More recently, a number of variables have been found to modulate the degree of displacement and contradict the literal explanation of internalised momentum as well as providing further evidence against an influence of eye-movements. For example, conceptual knowledge about a target object's typical movement can influence the magnitude of displacement. Reed and Vinson (1996) found differences in the Representational Momentum effect for two identical moving targets that only differed in the labels they had been assigned. An upwards moving target that was labelled as a rocket elicited more forward displacement than when it was labelled as a church. When the visual features of these stimuli were manipulated, to reflect a canonical rocket or a more ambiguous rocket for example, more forward displacement was found for prototypical self-propelled objects (Vinson & Reed, 2002). Moreover,

forward displacement is also modulated by expectations of a future change in movement direction (Johnston & Jones, 2006; Verfaillie & d'Ydewalle, 1991) and by beliefs about the source of the target's motion (for example, after being contacted by a moving object or "launcher", Hubbard & Favretto, 2003; Hubbard & Ruppel, 2002). Together, these findings suggest that the Representational Momentum effect not only results from extracting low-level motion information from the target's movement to anticipate its future position, but is also influenced by top-down information from prior knowledge or experience, in line with predictive processing theories (Clark, 2013; Friston, 2010).

Predictive processing models of perception argue that what we perceive results from a series of hypothesis testing and revision, predicting the incoming sensory information from their inferred cause and updating the inferred cause as the sensory information is received, until the most likely cause of the input is derived (Clark, 2013; Friston, 2010). These models can explain perceptual processing in sensory domains across the brain and the resulting distortions to perception that are often evidenced by perceptual illusions. For example, they can account for the illusory motion that is perceived between two visual stimuli when they are flashed in quick succession at different positions on the screen (apparent motion, Kolers, 2013; Wertheimer, 1912). This apparent motion effect is argued to reflect the prior hypothesis of a single moving stimulus generating a predicted motion path between the two locations. The perceptual system fills in the predicted but missing steps in-between as the input, albeit partially, confirms this hypothesis (Muckli et al., 2005; Sterzer et al., 2006). Consistent with a contribution of prior knowledge, this predicted path does not just reflect the shortest distance between two locations, but is influenced by prior knowledge about the moving target and its capabilities, such that

apparent motion for biological stimuli is perceived to curve around obstacles that are in the way (Chatterjee, Freyd, & Shiffrar, 1996; Shiffrar & Freyd, 1990,1993).

Predictive processing theories can explain the Representational Momentum effect in the same way that they can account for the influence of prior knowledge on illusory motion perception in apparent motion studies. The Representational Momentum effect can be theorised as illusory or apparent motion that continues once the moving target has disappeared, and the magnitude and direction of the distortion similarly reflects expectations not only from bottom-up information from the movement of the target, such as its speed or direction (Finke et al., 1986; Hubbard, 2005; Hubbard & Bharucha, 1988a), but also top-down knowledge about the typical movement of the target object (Reed & Vinson, 1996; Vinson & Reed, 2002), assumptions about the source of movement (Hubbard & Favretto, 2003; Hubbard & Ruppel, 2002), and even expected changes in movement direction (Johnston & Jones, 2006; Verfaillie & d'Ydewalle, 1991). Furthermore, a greater influence from top-down information is experienced when visual information from the target is ambiguous or degraded (Kok et al., 2013; Lages et al., 2017). Therefore, when presented with forward probes that match anticipated future motion, they are more readily perceived as being in the same position as the last seen image of the moving target before its sudden offset. Backward probes, on the other hand, reflect an unexpected position and would elicit salient prediction errors that are more easily detected, in line with prediction processing theories.

Recent evidence for the perceptual nature of the Representational Momentum effect has been revealed using neuroimaging methods. Functional magnetic resonance imaging (fMRI) studies have revealed that the areas in the brain responsible for

motion perception (V5/MT) are activated by static images with implied motion (Kourtzi & Kanwisher, 2000) and give rise to the Representational Momentum effect (Senior et al., 2000a). Furthermore, when transcranial magnetic stimulation (TMS) is applied to these areas, the Representational Momentum effect is eliminated (Senior, Ward, & David, 2002). More direct evidence for a top-down effect on perceptual regions comes from studies showing primary visual cortex (V1) activation during illusory perception in apparent motion tasks, in the precise retinotopic locations for the perceptual “filling-in” of the missing visual information (Muckli et al., 2005; Yantis & Nakama, 1998). This is supported by findings that show that after training participants with dynamic stimulus sets, only presenting the initial stages of these motions triggers V1 activity that resembles the perception of the full trajectory (Ekman, Kok, & de Lange, 2017). Prior information and knowledge therefore informs perceptual predictions of future motion, as revealed during the sudden offset of motion stimuli, as in the Representational Momentum effect, and result in the perceptual filling in of the missing input. Other investigations have revealed that perceptual expectations do not only act on the sudden absence of visual information, but facilitate direct changes to ongoing perception by sharpening the sensory input (Hammett, 1997; Kok, Jehee, & de Lange, 2012; Yon, Gilbert, Lange, & Press, 2018). The anticipation of future motion therefore acts directly on perceptual regions, either during the ongoing motion or after its sudden offset, and can influence the perceptual experience of observed motion, in line with predictive processing theories.

The findings discussed here therefore show that the Representational Momentum paradigm is an appropriate measure of predictive processing, providing a measure of expectation that not only draws upon bottom-up information provided by the

incoming visual stimuli, but is also influenced by top-down information from prior knowledge, reflected in changes to perception either during the ongoing observation or after the sudden offset of stimuli. By replacing the motion stimuli used in these paradigms with biological action, Representational Momentum presents as a novel tool to reveal the perceptual predictions people make about other people's actions and the consequences that these predictions have on social perception.

2.3 Representational Momentum to measure social prediction

In non-social perception, predictions of future states in Representational Momentum paradigms can be informed by variables such as properties of the target (Finke et al., 1986; Hubbard, 2005; Hubbard & Bharucha, 1988; Reed & Vinson, 1996; Vinson & Reed, 2002), inferred motion source (Hubbard & Favretto, 2003; Hubbard & Ruppel, 2002), and physical forces (Freyd & Finke, 1985; Hubbard, 1995b, 1998), for example. As described above, expectations for human action must additionally consider what is biomechanically possible (Chatterjee, Freyd, & Shiffrar, 1996; Shiffrar & Freyd, 1990,1993), to represent an accurate inference about predicted action. In addition, human action is less directly determined by physical variables such as momentum or friction. Instead, human actions are driven by the mental states of the actor, their goals and intentions, given the environmental context (Bach, Bayliss, & Tipper, 2011; Hamilton, 2009; Sebanz & Knoblich, 2009). Therefore, in order to more accurately predict the upcoming actions of others, one must be receptive to cues that signal their intention and integrate them into one's predictions.

There is some evidence of the use of Representational Momentum paradigms to measure such social prediction processes, but these investigations are not extensive

and the findings are questionable as to whether they reveal perceptual predictions born out of mental state attribution. Yoshikawa and Sato (2008) were one of the first to measure the Representational Momentum effect for dynamic facial expressions. When participants were shown images of faces that morphed from a neutral facial expression to one of the six basic emotions (Ekman, 1993), they reported seeing a more developed facial expression than that shown, providing evidence of a forward displacement towards more extreme expression. When the velocity of the morph was increased, these exaggerations were intensified. This provides initial evidence that the Representational Momentum paradigm can be used to capture predictions not only in the non-social domain, but also for human biological action. However, these findings are limited to low-level predictions based on the natural development of the observed faces and do not require higher-level top-down predictions about the actor's mental state.

When expressions of approach (joy and anger) and avoidance (fear and disgust) were measured against congruent or incongruent behaviours in terms of a rotating head towards or away from the observer, the final position of the head was judged to have rotated further when the expression matched the behaviour (i.e., further towards the participants for approach expressions and further away for avoidance expressions, Hudson & Jellema, 2011). These displacements demonstrate a Representational Momentum effect that draws not only upon predictions of continuing movement, but integrates top-down inferences about the actor's intention, as signalled by the match between their facial expression and the direction of their head turn. These findings therefore demonstrate the potential for the Representational Momentum paradigm to reveal the underlying goals and intentions that we attribute to others, in line with predictive processing theories.

Recent research has used the Representational Momentum paradigm to investigate whether high-level expectations about other people's actions do indeed induce perceptual changes to action observation. Hudson, Nicholson, Simpson, Ellis, and Bach (2016) presented participants with short video clips of an actor reaching towards or withdrawing away from an object. Before the onset of the action, participants instructed the actor by saying either "Take it!" or "Leave it!". Participants overestimated the last seen position of the hand in the direction of motion, and this forward bias was larger when the direction of the action followed the instruction. Thus, reaches towards an object appeared more exaggerated after instructing them to "Take it!" than to "Leave it!", and vice versa for observed withdrawals. This Representational Momentum effect was replicated when the actor stated their intention instead, by saying "I'll take it!" or "I'll leave it" before the onset of the action (Hudson, Nicholson, Ellis, & Bach, 2016), and was modulated by the likelihood that the actor would do as he says (25% or 75% of the time, Hudson, Bach, & Nicholson, 2018). These findings therefore reveal that high-level information about the goals and intentions of other's actions is used to generate predictions of their upcoming behaviour, that these predictions create distortions to the perceptual representation of the action, and that these social biases can be captured within a Representational Momentum paradigm.

Together, therefore, the Representational Momentum effect presents a unique measure of predictive motion perception and has been used extensively to reveal how perceptual predictions about the future path of a moving target can distort the perceptual representation of the stimulus (Freyd & Finke, 1984; Hubbard, 2005, 2017). These predictions are not only informed by low-level motion cues from the observed stimulus, but are also influenced by higher-level expectations that consider

the features of the target (Finke et al., 1986; Hubbard, 2005; Hubbard & Bharucha, 1988; Reed & Vinson, 1996; Vinson & Reed, 2002), the environment (Johnston & Jones, 2006; Verfaillie & d'Ydewalle, 1991), and the causes of motion (Hubbard & Favretto, 2003; Hubbard & Ruppel, 2002), for example. More recently, the Representational Momentum effect has been employed as a novel method to measure social predictions, showing modulations to the predictive effects based on inferences about the mental state of the actor (Hudson et al., 2018; Hudson & Jellema, 2011; Hudson, Liu, & Jellema, 2009; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016). These findings have shown the capabilities of Representational Momentum paradigms to capture the high-level expectations that humans have about other people's upcoming behaviours, informed by inferences about their goals and intentions, and the influence these predictions have on social perception. While these prior studies demonstrate initial evidence for comparable predictive processing mechanisms between social perception and non-social perception, offering a revolutionary perspective on how we understand the actions of others, the findings are limited and so many questions about the underlying mechanisms of action understanding remain unanswered. The Representational Momentum paradigm therefore offers novel method to conduct deeper investigations into predictive processing theories of social perception. The experiments in this thesis will use this paradigm to reveal how these predictions of others' upcoming behaviour are generated, on what cues they rely, and how these predictions alter the perceptual representation of others' action.

3 Chapter Three - Assumptions of efficient action

The initial studies aimed to resolve which information is used to inform action predictions and what impact these predictions have on social perception, therefore providing new avenues to understand how action expectations can shape our understanding of other people's actions. These studies use the Representational Momentum paradigm (Freyd & Finke, 1984; Hubbard, 2005) to test whether assumptions of action efficiency (Csibra & Gergely, 2013; Dennett, 1987; Gergely & Csibra, 2003) inform expectations of others upcoming actions. They test whether these predictions specifically concern the expected action kinematics in a perceptual format, how these predictions bias the actions' perceptual representation, and whether these perceptual biases are elicited spontaneously.

The experiments in this chapter were published in Proceedings of the Royal Society: Biological Sciences and are presented in their published format (green copy).

3.1 Perceptual Teleology: Expectations of Action Efficiency Bias Social Perception.

Matthew Hudson*, **Katrina L. McDonough***, Rhys Edwards, Patric Bach*
***authors contributed equally**

Abstract

Primates interpret conspecific behaviour as goal-directed and expect others to achieve goals by the most efficient means possible. While this teleological stance is prominent in evolutionary and developmental theories of social cognition, little is known about the underlying mechanisms. In predictive models of social cognition, a perceptual prediction of an ideal efficient trajectory would be generated from prior

knowledge against which the observed action is evaluated, distorting the perception of unexpected inefficient actions. To test this, participants observed an actor reach for an object with a straight or arched trajectory on a touch screen. The actions were made efficient or inefficient by adding or removing an obstructing object. The action disappeared mid-trajectory and participants touched the last seen screen position of the hand. Judgments of inefficient actions were biased toward the efficient prediction (straight trajectories upward to avoid the obstruction, arched trajectories downward towards the target). These corrections increased when the obstruction's presence/absence was explicitly acknowledged, and when the efficient trajectory was explicitly predicted. The teleological stance is at least partly perceptual, providing an ideal reference trajectory against which actual behaviour is evaluated.

3.2 Experiments 1a - c

Human and non-human primates take the "intentional stance" when watching conspecifics (Dennett, 1987), interpreting their behaviour as purposeful and goal directed (Baillargeon, Scott & Bian, 2016; Baker, Saxe & Tenenbaum, 2009; Csibra & Gergely, 2007; Gergely & Csibra, 2003). Crucial to this is the understanding that others' actions are, from the outset, optimised to achieve their goals in the most efficient and rational way, minimising time and energy expenditure given the environmental constraints. Both human infants and macaque monkeys, for example, show surprise when intentional agents do not attempt to avoid an obstacle, or take an unnecessary long way to reach their goal (Gergely & Csibra, 2003; Rochat, Serra, Fadiga & Gallese, 2008). This simple efficient action heuristic provides a foundation for the development of sophisticated capacities for mentalizing and theory of mind in adult humans (e.g., Gergely & Csibra, 2003; Wellman & Brandone, 2009). For

example, seeing a seemingly inefficient action (e.g. a reach straight for an object despite an obstacle in the way) can prompt the insight that others act according to beliefs that can differ from one's own (i.e. they may not have seen the obstacle). Indeed, seeing such actions captures attention (Vivanti et al., 2011), and alters activity in brain areas implicated in action perception and mentalizing (e.g. Desmet & Brass, 2015; Marsh, Mullett, Ropar & Hamilton, 2014).

Yet, despite the crucial role of teleological/intentional reasoning in human and animal social cognition, little is known about the underlying processes. The currently dominant view sees social perception as a bottom-up "resonance" of one's own motor apparatus with others' actions, which allows the associated goals and internal states (sensations, emotions) to be derived (e.g., Iacoboni, 2009; Rizzolatti & Sinigaglia, 2010). Action efficiency would, in such a model, be conceptualised as a post-hoc motoric signal of effort or energy expenditure, which can be compared with a reference value for this type of action (e.g. Jara-Ettinger, Gwean, Tenenbaum and Schulz, 2015). However, such models are challenged by findings that children make efficiency judgments for movements of biomechanically impossible actions for which motor resonance is unlikely (Southgate, Johnson & Csibra, 2008), that they can process efficiency before acquiring competence in the seen actions (e.g., Gredeback & Melinder, 2010; Sodian, Schoeppner & Metz, 2004), or that, in adults, eye-movements indicate expectations of efficient action before action onset, when such kinematic information is not yet available (Fischer, Prinz & Lotz, 2008).

An alternative is that teleological reasoning might not emerge from a "late" motoric signal, but from earlier perceptual signals. Recent predictive coding frameworks argue that perception in general – and social perception in particular – is informed by prior expectations, derived from one's knowledge about the world and

other people, and that these expectations guide processing of the perceptual input (Bach, Nicholson & Hudson, 2014; 2015; Bach & Schenke, 2017; Csibra, 2008; Kilner, Friston & Frith, 2007[ab]). Predictive influences have been demonstrated in a diverse range of perceptual abilities including the perception of "true" colour from surrounding illumination (Bloj, Kersten & Hurlbert, 1999), anticipated effects of physical dynamics on motion perception (Sotiropoulos, Seitz & Series, 2011), and 3D concave/convexity from the presumed location of light sources (Adams, Graf & Ernst, 2004). In a similar way, the environment provides all the necessary information to generate an ideal reference trajectory that a fully rational, intentional actor would take to achieve their goal (i.e. location of goal objects and possible obstructions), and which would provide a comparison to immediately flag observed actions as being efficient or not, confirming prior attributions of goals and intentionality.

Here, we provide a first test of (1) whether human observers make such predictions of how rational actors who are aware of all environmental constraints efficiently traverse the given action space, (2) whether these predictions are realised in a perceptual format that can serve as a reference image for the observed action and (3) which bias its perceptual representation. We rely on the well-established phenomenon that when a moving stimulus suddenly disappears, participants' estimations of its last seen position show robust distortions towards the expected path (i.e. Representational Momentum, Hubbard, 2005), in line with the notion that the considerable uncertainty during motion perception is sharpened by top-down information (e.g., Hammett, 1997; Bex, Edgar & Smith, 1995), or that predicted paths are perceptually "filled in" after the sudden offset (Ekman, Kok & de Lange, 2017). Importantly, these distortions rely on changes to lower-level visual representations

(e.g. Senior, 2000; Senior, Ward & David, 2002), occur even when participants are warned against them (Courtney & Hubbard, 2008; Ruppel, Fleming & Hubbard, 2009) and integrate higher level information such as the physical forces acting on the objects (e.g., momentum, friction, gravity, for a review, see Hubbard, 2005) or prior action expectations (Hudson, Nicholson, Ellis & Bach, 2016; Hudson, Nicholson, Simpson, Ellis & Bach, 2016; Hudson, Bach & Nicholson, 2017).

Here, we use this paradigm to reveal the expectations of efficient action that guide the perception of others' actions. In three studies, participants watched an actor reach towards an object. The action disappeared mid-trajectory, and participants indicated the perceived disappearance point on a touch screen. In two conditions, the actions were efficient, showing a reach either straight towards the object or arched over an obstacle placed in between. In two other conditions, the actions were made inefficient by either adding an obstacle to the path of the straight reaches (such that the actor would knock into the obstacle), or removing the obstacle for the arched reaches (such that the actor reached over empty space). If others' behaviour is perceived relative to what would be expected under the implicit assumption of efficient action, then the perceived kinematics should be displaced along the trajectory that an intentional, rational actor might take. Unexpected inefficient actions should be "corrected" toward the predicted efficient action trajectory: straight reaches would be perceived upward if approaching an obstacle where an avoidance movement would be predicted, whilst an arched reach would be displaced downwards if made over empty space as this energy expenditure is unnecessary. Moreover, such distortions should be observed spontaneously when participants passively observe these actions, but should increase the more the environmental constraints and the behaviour of a rational actor is made explicit. As

an additional between-subjects task manipulation, we therefore varied whether the actions were viewed under no additional instructions (No Task), or whether they were asked to report “yes” or “no” in response to the presence of an obstacle prior to action onset (Report Obstacle) or to predict whether a rational actor would ideally have to reach “straight” or “over” an obstacle before the action started (Predict Trajectory).

3.2.1 Method

3.2.1.1 Participants

Eighty-five participants took part (mean age=24 years, SD=7.7, 62 females, No Task: *n*=30, Report Obstacle: *n*=27, Predict Trajectory: *n*=28). Seven additional participants were excluded due to performance (see Results). All participants were right handed, had normal/corrected vision, were recruited from Plymouth University and wider community, and received course credit or payment. The study received ethical approval from the University of Plymouth’s ethics board, in accordance with those of the ESRC and the Declaration of Helsinki. A priori power analyses of previous experiments investigating similar effects with the same method (Hudson et al., 2017, Experiment 3) revealed that a sample size of 14 is required to achieve power of 0.95.

3.2.1.2 Apparatus

Stimuli were filmed with a Sony HD video camera at 50fps with a widescreen aspect ratio (16:9) and a resolution of 1920 X 1080 (2.1 megapixel), and edited with Adobe Photoshop. The experiment was delivered using Presentation (NeuroBS) via

a NEC Multisync P221w LCD touch screen monitor (1680 X 1050). Verbal responses for the Report Obstacle and Predict Trajectory conditions were recorded using Presentation's sound threshold logic via a Logitech PC120 combined microphone and headphone set.

3.2.1.3 Stimuli

Example stimuli can be seen in Figure 3-1A. Videos were filmed of an arm starting in a rest position at the right of the screen and reaching to grasp a target object on the left (either an apple, bottle, crisps, glue stick, or stapler). In the original set of videos, the actor´s reach was either (1) unobstructed and the trajectory of the arm was straight toward the target object (Straight/Efficient), (2) obstructed by one of 4 objects (iPad, lamp, pencil holder, or photo-frame), and the trajectory of the arm was arched over the obstruction (Arched/Efficient). From each video, 19 frames were extracted for the experimental stimuli, beginning with the onset of movement (frame 1) to mid-way through the action (frame 19). Inefficient action sequences were created by digitally removing the obstructing objects in the Arched/Efficient videos (Arched/Inefficient). For each of the Straight/Efficient actions, a new set of videos were created by adding each of the obstructing objects to show the actor was reaching straight for the target despite the obstruction (Straight/Inefficient). This created a set of inefficient actions that were identical to the efficient actions in terms of movement kinematics, and differed only by the presence/absence of the obstructing object. Finally, for each action a single frame was created in which the hand was digitally removed. This served as a response stimulus at the end of each trial where participants estimated the disappearance point of the action.

3.2.1.4 Procedure

Participants completed two blocks of 80 trials in which each combination of action trajectory (straight, arched) and efficiency (efficient, inefficient) was represented by 20 trials. Participants were instructed that, on each trial, they would see an actor reach from the right of the screen for a target object on the far left, but that sometimes there would be a second object in between.

An example trial sequence can be seen in Figure 3-1B. At the start of each trial, participants were instructed to “Hold the spacebar” and to keep it depressed to prevent them from tracking the observed action with their finger to improve performance. They then saw the first frame of the action sequence as a static image. In the No Task condition, the action sequence began after a randomly chosen delay of between 1000ms or 3000ms. In the Report Obstacle condition and Predict Trajectory conditions, it began 1000ms after the participant’s verbal response had been detected. In the Report Obstacle condition, participants said “No” if there was no obstruction and “Yes” if there was an obstruction. In the Predict Trajectory condition, participants said “Straight” if there was no obstruction, and “Over” if there was an obstruction. The action depicted the frame order progressing at 3 frame intervals for a total sequence of between 4 and 7 frames for 80ms each (e.g. frames 1-4-7-10-13-16-19). Starting frames and sequence length were randomly chosen on each trial to prevent memorisation of the final position from the starting frame. The final frame was immediately replaced by the response stimulus, creating the impression that the hand simply disappeared from the scene. Participants released the spacebar and, with their right hand, touched the screen where they thought the final seen position of the tip of the index finger was. As soon as a response was registered, the next trial began.

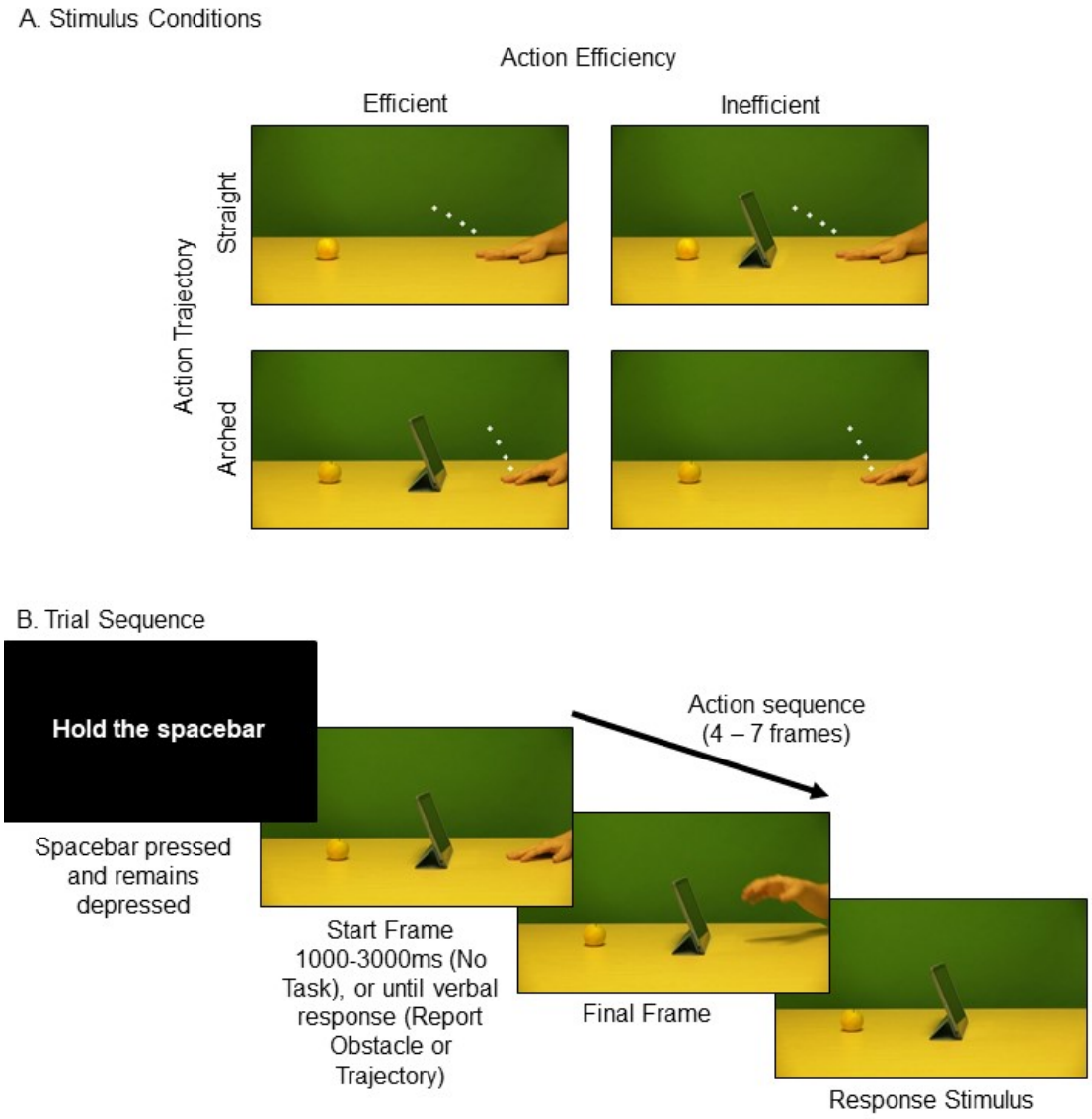


Figure 3-1. Experiments 1a-c. Stimulus conditions and trial sequence.

The stimulus conditions are depicted in Panel A. The Action Trajectory was either straight (top row) or arched over (bottom row). The presence or absence of an obstructing object made the action trajectory either efficient (left column) or inefficient (right column). In all examples, the hand is in the initial start position, and the white markers depict the final four frames of the trajectory of the index finger tip. The action sequence disappeared at one of these four points. An example trial sequence is depicted in Panel B, depicting an efficient arched trajectory over an obstruction.

3.2.2 Results

Participants were excluded if the distance between the real and selected positions exceeded 3SD of the sample mean (mean =49.2 pixels, *SD*=12.4, no

exclusions), or if the correlation between the real and selected positions was more than 3SD below the median r value (X axis: median =.914, *SD* = .055; Y axis: median =.901, *SD* = .077, 4 participants excluded). For each participant, individual trials were excluded if the response procedure was incorrect (spacebar released before action offset, 4.1%), or if response initiation or execution times were less than 200ms or more than 3SD above the sample mean (5.1%, Initiation: mean =443.5ms, *SD*=84.4; Execution: mean =817.2ms, *SD*=230.4). Three additional participants were excluded for having an excessive number of trial exclusions (> 50%).

The real final screen coordinate of the tip of the index finger was subtracted from participants' selected screen coordinate on each trial. Analysis was conducted on this residual localisation error, which provided a directional measure of how far, in pixels, participant's responses were displaced along the X and Y axis. An accurate response would produce a value of 0 on both axes. On the X axis, positive values denote a rightward displacement (against the direction of motion) and negative values a leftward displacement. On the Y axis positive and negative values denote upward and down displacements respectively.

Overall, there was a significant leftward bias (X axis: mean =-8.4px, *SD*=19.2, *t*(84)=-4.01, *p*<.001, *d*=.61, *95% CI* [-4.3,-12.5]), and a significant downward bias (Y axis: mean =-15.1px, *SD*=15.0, *t*(84)=-9.27, *p*<.001, *d*=1.35, *95% CI* [-11.9,-18.3]). The differences along the X and Y axis for each experimental condition across all tasks and for each task individually can be seen in Figure 3-2: A-D. These difference values were entered into a 2X2X3 mixed measures ANOVA for the X and Y axis separately, with Trajectory (arched, straight) and Efficiency (efficient, inefficient) as within-subjects factors, and Task (No Task, Report Obstacle, Predict Trajectory) as a between-subjects factor.

3.2.2.1 Y Axis

The main prediction is that perceptual judgments of inefficient actions would be displaced towards the expected trajectory, that is, downwards for inefficient arched trajectories and upwards for inefficient straight trajectories. Indeed, the analysis revealed a main effect of Efficiency ($F(1,82)=12.04$, $p=.001$, $\eta_p^2=.128$) that was qualified by an interaction of Efficiency and Trajectory ($F(1,82)=136.2$, $p<.001$, $\eta_p^2=.624$). As predicted, inefficient arched trajectories (-19.7px) were displaced below efficient arched trajectories (-11.0px, $t(84)=-9.33$, $p<.001$, $d=.47$), and inefficient straight actions (-12.0px) were displaced above efficient straight actions (-17.7px, $t(84)=8.51$, $p<.001$, $d=.44$), despite the actual hand disappearance points being identical within each trajectory. Importantly, there was a three-way interaction between Trajectory, Efficiency, and Task ($F(2,82)=10.6$, $p<.001$, $\eta_p^2=.205$). The interaction effect was re-quantified as a single value for each participant ([Arched/Efficient – Arched/Inefficient] – [Straight/Efficient – Straight/Inefficient]) to reveal the total amount in pixels by which inefficient actions were corrected toward a more efficient trajectory for each task (see Figure 3-2: E). Between subjects t-tests on this interaction value – mathematically equivalent to the pairwise 3-way interactions of Trajectory, Efficiency and Task – show that the interaction was marginally larger in the Predict Trajectory condition than the Report Obstacle condition ($t(53)=-1.95$, $p=.057$, $d=.53$), which in turn was significantly larger than in the No Task condition ($t(55)=-2.81$, $p=.007$, $d=.73$). Demonstrating the robustness of the interaction, exploratory analysis showed that the interaction of Trajectory and Efficiency was evident in all conditions with a corrected alpha level of $p = .017$ (No Task: $F(1,29)=35.3$, $p<.001$, $\eta_p^2=.549$; Report Obstacle: $F(1,26)=44.5$, $p<.001$,

η_p^2=.631; Predict Trajectory: F(1,27)=57.7, p<.001, η_p^2=.681, Figure 2 A-C). There were no further main effects or interactions.

3.2.2.2 X Axis

We did not have specific predictions about how action rationality would affect perceptual displacements on the X Axis and the analysis indeed did not reveal either a main effect of Efficiency (F(1,82)=.837, p=.363, η_p^2=.01) nor an interaction of Efficiency and Trajectory (F(1,82)=1.39, p=.242, η_p^2=.017). We report the remaining effects of no interest below but due to alpha inflation of unpredicted effects in an ANOVA (Cramer et al., 2014) they should be considered exploratory and interpreted with caution. A main effect of Task (F(2,82)=8.81, p<.001, η_p^2=.177) revealed a general leftward displacement in the No Task (-16.8px, t(29)=-5.14, p<.001, d=1.35) and Predict Trajectory conditions (-10.0px, t(27)=-3.43, p=.002. d=.86), but not in the Report Obstacle condition (9px, t(26)=.720, p=.478, d=.19). A main effect of Trajectory (F(1,82)=1231.4, p<.001, η_p^2=.938) showed a leftward displacement for arched trajectories (-24.4px, t(84)=-11.0, p<.001, d=1.7) and a rightward displacement for straight trajectories (7.7px, t(84)=3.74, p<.001, d=.53), most likely reflecting the further right displaced centre of gravity of the straight arm configurations (Coren & Hoeing, 1972; see also Hudson et al., 2017). Finally, an interaction between Trajectory and Task (F(1,82)=9.88, p<.001, η_p^2=.194) revealed that the Trajectory effect was larger in the No Task than in the Report Obstacle condition (t(55)=4.15, p<.001, d=1.1) and Predict Trajectory conditions (t(56)=3.25, p=.002, d=.86), which did not differ from each other (t(53)=1.18, p=.243, d=.32). There were no further interactions (all p's>.351).

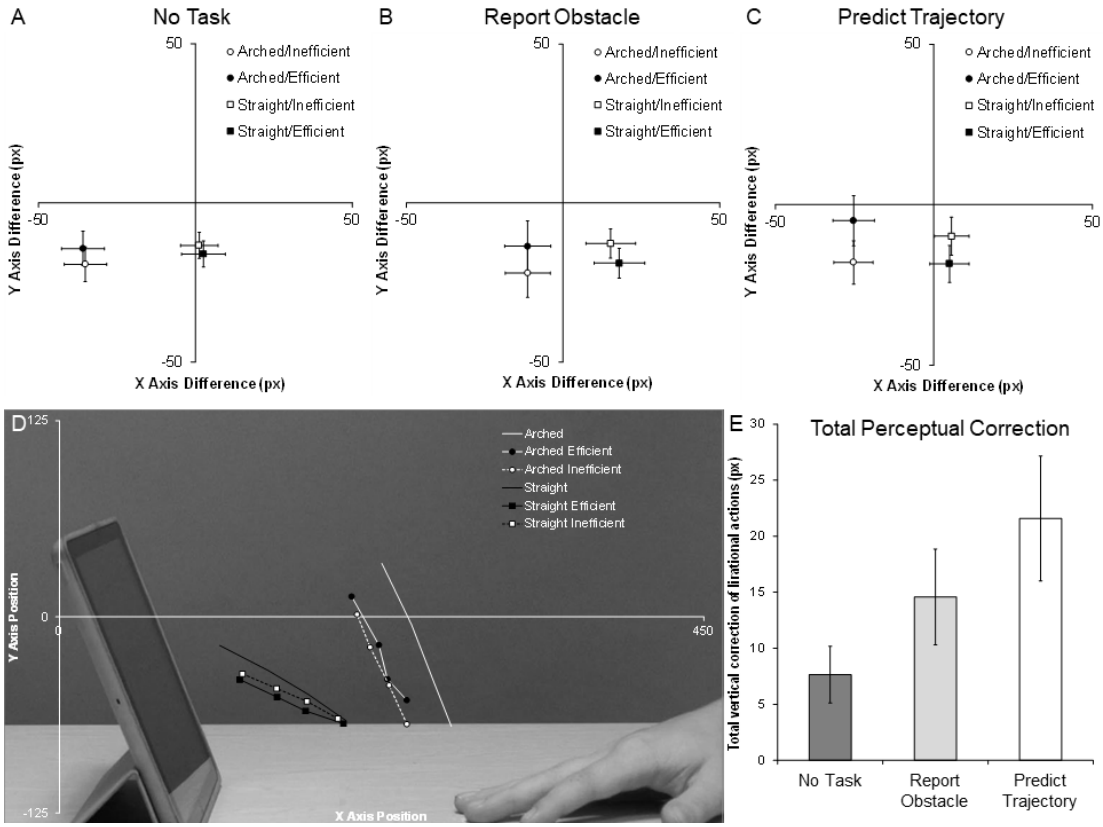


Figure 3-2. Experiments 1a-c. Results

The Trajectory X Efficiency interactions for each Task condition are depicted on the top row (A: No Task; B: Report Obstacle; C: Predict Trajectory). The difference between the real final position and the selected final position is plotted for the X axis and Y axis. The centre of each plot represents the real final position on any given trial (0px difference on each axis). Panel D provides a descriptive representation of the data in real screen coordinates (collapsed across task conditions). The sold lines represent the mean real final position of the Arched (white) and Straight (black) trajectories for the four possible disappearance points. The selected screen coordinates for each Trajectory are plotted for the efficient (filled line) and inefficient (dashed line) conditions. The data are placed over a spatially aligned backdrop of a representative stimulus image of the action start point with an obstructing object to provide a reference of how the data relate to the stimuli. Panel E depicts a comparison of the size of the Y axis interaction in pixels, equivalent to the total amount by which inefficient actions were corrected towards a more efficient trajectory. Error bars depict 95% confidence intervals.

3.2.3 Discussion

The present study showed for the first time that the teleological interpretation humans have of others' behaviour is perceptually instantiated and provides a visual

reference signal for an expected “ideal” trajectory during action observation. Participants watched a hand reach for objects with either efficient or inefficient kinematics and reported its last position after it had suddenly disappeared. Across several samples, perceptual reports were consistently biased towards the ideal reference kinematics. Straight reaches were reported higher if there was an obstacle in the way, as if lifted to avoid it. Conversely, reaches with a high arched trajectory were reported lower if the path was clear. These biases were evident automatically, but became more pronounced when observers explicitly processed the potential obstacles that could constrain the action and particularly when they predicted the most efficient action kinematics through the scene.

Together, these results reveal that, during social perception, the principle of efficient action provides a similar perceptual reference signal as the assumption that light comes from above (Adams et al., 2004) or that gravity pulls objects downwards (Hubbard & Bharucha, 1988), constantly pushing the perceptual representation of inefficient actions towards a more rational path. The resulting biases in perceptual judgments cannot be accounted for by an abstract awareness of the action’s goal, such as when eye movements jump towards an action’s target (Eshuis, Coventry & Vulchanova, 2009) or perceptual judgments are biased towards them (Hudson et al., 2016[ab]; 2017). Instead, they reveal concrete expectations of the specific trajectory that the action will take through the scene. Moreover, making this awareness explicit prior to action observation increased the perceptual bias. Action efficiency is therefore not only evaluated after an action has been completed (Gergely & Csibra, 2003; Rochat et al., 2008; Scott & Baillargeon, 2013; Vivanti et al, 2011), but constantly updated, at every step in the trajectory, by predictions that can be derived from contextual cues prior to the motor behaviour.

Our results support predictive coding frameworks of social perception (Bach et al., 2014; 2015; 2017; Csibra, 2008; Kilner et al., 2007[ab]), which argue that social perception, like perception in general, is hypothesis driven and guided by top-down expectations. In such models, observers constantly test their inferences about others' goals and beliefs by predicting how they would behave, and matching this prediction to – and integrating it with –the actual perceptual input. In such a view, predictions of efficient action can contribute to the perceptual sharpening of the visual uncertainty during action perception (i.e. motion blurring, Hammett, 1997), or after its offset, constantly biasing perception toward the expected avoidance or straightening movement, with the amount of bias constrained by the visual uncertainty. In addition, they allow humans to rapidly confirm the intentionality of other's behaviour, only requiring a match of the observed actions to the ideal kinematics that would follow from these goals. Prediction errors, in contrast, would signal inefficient actions, triggering more sophisticated mentalising processes to re-evaluate the actor's goal or how their beliefs may differ from one's own (Desmet & Brass, 2015; Gergely & Csibra, 2003; Marsh et al., 2014). In this way, the relatively simple perceptual process of prediction and prediction error would not only support perception, but also provide a foundation for higher level judgments about others' beliefs or intentions, even in cases in which motor experience is unlikely (Gredeback & Melinder, 2010; Sodian et al., 2004; Southgate et al., 2008).

In humans, sensitivity to kinematic efficiency emerges early in development. It is present in other primates (Rochat et al., 2008) and is also spared in individuals with autism spectrum conditions, for whom only more advanced mental state reasoning prove problematic (Marsh, Pearson, Ropar & Hamilton, 2015; Vivanti et al, 2011). The key leap to the sophisticated socio-cognitive abilities of humans may

therefore lie in the abstraction of these content-based representations of goals, environment, and action, to higher-order representations of desires, beliefs, and intentions, respectively (Gergely & Csibra, 2003). For example, it is clear that human infants from the age of four onwards are able to predict what others will do not based on the actual environmental constraints (toy is in box A), but in terms of what the actor believes the state of the environment to be (they believe it is in box B), suggesting that the bottleneck emerges at this later state that requires sophisticated coordination of representation, such as others' beliefs that differ from one's own beliefs or objective reality (Samson, Apperly, Kathirgamanathan & Humphreys, 2005; Santiesteban, White, Cook, Gilbert, Heyes & Bird, 2012; Steinbeis, 2016).

3.2.4 Conclusions

The principle of efficient action allows observers to perceive others' actions relative to ideal reference actions, thereby confirming prior goal attributions or revising them in case of a conflict. Such perceptual mechanisms for rationality perception support rapid attribution of intentionality and facilitate the perception of others' behaviour and our interactions with them. The burden of social cognition is placed on mechanisms that account for unexpected behaviour through a re-evaluation of their beliefs, desire, and intentions, so that our model of the social world can be refined, and predictions of other's behaviour can be made more accurately.

3.2.5 References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7,* 1057-1058. doi:10.1038/nn1312

Andrade, J., Kemps, E., Werniers, Y., May, J., & Szmalec, A. (2002). Insensitivity of visual short-term memory to irrelevant visual information. *The Quarterly Journal of Experimental Psychology: Section A*, *55*, 753-774. doi:10.1080/02724980143000541

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: How objects guide action understanding and prediction. *Frontiers in Human Neuroscience,* 8:254. doi:10.3389/fnhum.2014.00254

Bach, P., Nicholson, T., & Hudson, M. (2015). Pattern completion does not negate matching: A response to Uithol and Maranesi. *Frontiers in Human Neuroscience.* 9:685. doi:10.3389/fnhum.2015.00685

Bach, P., & Schenke, K. C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass*, *11*. doi:10.1111/spc3.12312

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological reasoning in infancy. *Annual Review of Psychology, 67,* 159-186. doi:10.1146/annurev-psych-010213-115033

Baker, C.L., Saxe, R., & Tenenbaum, J.B. (2009). Action understanding as inverse planning. *Cognition, 113,* 329-349. doi:10.1016/j.cognition.2009.07.005

Bex, P. J., Edgar, G. K., & Smith, A. T. (1995). Sharpening of drifting, blurred images. *Vision Research, 35*, 2539-2546. doi:10.1016/0042-6989(95)00060-D

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature, 402*, 877-879. doi:10.1038/47245

Borst, G., Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2012). Representations in mental imagery and working memory: Evidence from different types of visual masks. *Memory & Cognition*, *40*, 204-217. doi:10.3758/s13421-011-0143-7

Breitmeyer, B., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (No. 41). Oxford University Press.

Coren, S., & Hoenig, P. (1972). Effect of non-target stimuli upon length of voluntary saccades. *Perceptual and Motor Skills*, *34*, 499-508. doi:10.2466/pms.1972.34.2.499

Csibra, G. (2008). Action mirroring and action understanding: An alternative account. *Sensorymotor Foundations of Higher Cognition. Attention and Performance XXII,* 435-459. doi:10.1093/acprof:oso/9780199231447.003.0020

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals: Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica, 124*, 60–78. doi:10.1016/j.actpsy.2006.09.007

Courtney, J. R., & Hubbard, T. L. (2008). Spatial memory and explicit knowledge: An effect of instruction on representational momentum. *The Quarterly Journal of Experimental Psychology, 61*, 1778-1784. doi:10.1080/17470210802194217

Cramer, A. O., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P., ... & Wagenmakers, E. J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review, 23,* 640-647. doi:10.3758/s13423-015-0913-5

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Desmet, C. & Brass, M. (2015). Observing accidental and intentional unusual actions is associated with different subregions of the medial frontal cortex. *NeuroImage, 122,* 195-202. doi:10.1016/j.neuroimage.2015.08.018.

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*, 15276. doi:10.1038/ncomms15276

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009). Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychological Science, 20,* 438-440. doi:10.1111/j.1467-9280.2009.02317.x

Firestone, C & Scholl, B.J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences, 39,* e229. doi:10.1017/S0140525X15000965

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, *19*, 1488-1497. doi:10.1162/jocn.2007.19.9.1488

Firestone, C & Scholl, B.J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences, 39,* e229. doi: 10.1017/S0140525X15000965

Fischer, M. H., Prinz, J., & Lotz, K. (2008). Grasp cueing shows obligatory attention to action goals. *The Quarterly Journal of Experimental Psychology, 61,* 860-868. doi:10.1080/17470210701623738

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in Cognitive Sciences, 7*, 287-292. doi:10.1016/S1364-6613(03)00128-1

Gredeback, G., & Melinder, A. (2010). Infants understanding of everyday social interactions: A dual process account. *Cognition, 114,* 197-206. doi:10.1016/j.cognition.2009.09.004

Hammett, S. T. (1997). Motion blur and motion sharpening in the human visual system. *Vision Research, 37*, 2505-2510. doi:10.1016/S0042-6989(97)00059-X

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12*, 822-851. doi:10.3758/BF03196775

Hubbard, T.L., & Bharucha, J.J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics, 44,* 211–221. doi: 10.3758/BF03206290

Hudson, M., Bach, P., & Nicholson, T. (2017). You said you would! The predictability of other's behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance, 44,* 320-335. doi:10.1037/xhp0000451

Hudson, M., Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition, 146*, 245-250. doi:10.1016/j.cognition.2015.09.021

Hudson, M., Nicholson, T., Simpson, W. A., Ellis, R., & Bach, P. (2016). One step ahead: The perceived kinematics of others' actions are biased toward expected goals. *Journal of Experimental Psychology: General, 145*, 1-7. doi:10.1037/xge0000126

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology, 60,* 653-670. doi:10.1146/annurev.psych.60.110707.163604

Ianì, F., Mazzoni, G., & Bucciarelli, M. (2018). The role of kinematic mental simulation in creating false memories. *Journal of Cognitive Psychology*, 1-15. doi:10.1080/20445911.2018.1426588

Jara-Ettinger, J., Gweon, H., Tenenbaum, J.B., & Schulz, L.E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition, 140,* 14-23. doi:10.1016/j.cognition.2015.03.006

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007[a]). The mirror-neuron system: A Bayesian perspective. *NeuroReport, 18*, 619–623. doi:10.1097/WNR.0b013e3281139ed0 00001756-200704160-00018

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007[b]). Predictive coding: An account of the mirror neuron system. *Cognitive Processing, 8*, 159-166. doi:10.1007/s10339-007-0170-2

Kinsbourne, M., & Warrington, E. K. (1962). The effect of an after-coming random pattern on the perception of brief visual stimuli. *The Quarterly Journal of Experimental Psychology, 14*, 223-234. doi:10.1080/17470216208416540

Lamme, V. A., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience, 14*, 1044-1053. doi:10.1162/089892902320474490

Marsh, L. E., Mullett, T.E., Ropar, D., & Hamilton, A.F. de C. (2014). Responses to irrational actions in action observation and mentalising networks of the human brain. *Neuroimage, 103,* 81-90. doi:10.1016/j.neuroimage.2014.09.020

Marsh, L.E., Pearson, A., Ropar, D., & Hamilton, A.F. de C. (2015). Predictive gaze during observation of irrational actions in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders, 45,* 245-261. doi:10.1007/s10803-014-2215-6

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology, 3*, e265. doi:10.1371/journal.pbio.0030265

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience, 11,* 264–274. doi:10.1038/nrn2805.

Rochat, M. J., Serra, E., Fadiga, L., & Gallese, V. (2008). The evolution of social cognition: Goal familiarity shapes monkeys' action understanding. *Current Biology, 18*, 227-232. doi:10.1016/j.cub.2007.12.021

Ruppel, S. E., Fleming, C. N., & Hubbard, T. L. (2009). Representational momentum is not (totally) impervious to error feedback. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 63*, 49-58. doi:10.1037/a0013980

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain, 128,* 1102-1111. doi:10.1093/brain/awh464

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to theory of mind. *Cognition, 122,* 228-235. doi:10.1016/j.cognition.2011.11.004

Scott, R.M., & Baillargeon, R. (2013). Do infants really expect agents to act efficiently? A critical test of the rationality principle. *Psychological Science, 24,* 466-474. doi: 10.1177/0956797612457395

Senior, C., Barnes, J., Giampietroc, V., Simmons, A., Bullmore, E. T., Brammer, M., & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or 'representational momentum'. *Current Biology, 10*, 16-22. doi:10.1016/S0960-9822(99)00259-6

Senior, C., Ward, J., & David, A. S. (2002). Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Visual Cognition, 9*, 81-92. doi:10.1080/13506280143000331

Sodian, B. Schoeppner, B., & Metz, U. (2004). Do infants apply the principle of rational action to human agents? *Infant Behavior and Development, 27,* 31–41. doi:10.1016/j.infbeh.2003.05.006

Sotiropoulos, G., Seitz, A.R. & Series, P. (2011). Changing expectations about speed alters perceived motion direction. *Current Biology, 21*, R883-R884. doi:10.1016/j.cub.2011.09.013

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition, 107*, 1059-1069. doi:10.1016/j.cognition.2007.10.002

Steinbeis, N. (2016). The role of self–other distinction in understanding others' mental and emotional states: Neurocognitive mechanisms in children and adults. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371,* 20150074. doi:10.1098/rstb.2015.0074

Tapia, E., & Beck, D. M. (2014). Probing feedforward and feedback contributions to awareness with visual masking and transcranial magnetic stimulation. *Frontiers in Psychology, 5*, 1173. doi:10.3389/fpsyg.2014.01173

Vivanti, G., McCormick, C., Young, G.S., Abucayan, F., Hatt, N., Nadig, A., Ozonoff, S., & Rogers, S.J. (2011). Intact and impaired mechanisms of action

understanding in autism. *Developmental Psychology, 47,* 841-856. doi:10.1037/a0023105.

Wellman, H. M., & Brandone, A. C. (2009). Early intention understandings that are common to primates predict children's later theory of mind. *Current Opinion in Neurobiology, 19,* 57-62. doi:10.1016/j.conb.2009.02.004

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience, 1*, 508-512. doi:10.1038/2226

4 Chapter Four: Perceptual predictions

The experiments in Chapter Three provide the first evidence that, during action observation, the principle of efficient action provides a perceptual reference signal for the specific trajectory a rational actor would make through a given scene, influencing action perception. Despite utilising a Representational Momentum paradigm (Freyd & Finke, 1984; Hubbard, 2005) that reliably measures predictive biases in motion perception, there still remains the possibility that the effects found could be influenced by changes outside of the perceptual system (Kerzel, 2005). For example, it is possible that participants merely misremembered the last final position of the actor's hand, or that the touch screen response measure introduced the potential for motor-related effects (Firestone & Scholl, 2015; Kerzel, 2005).

Experiment 2a tested whether the perceptual shifts towards efficient actions can be observed in a psychophysical task without motor or working memory components (Kerzel, 2005). This experiment replaced touch screen judgements with probe judgements directly after action offset (250 ms.) to rule out contributions from the action's representation in later working memory or motor control stages, and to instead reveal a contribution to immediate perceptual processing, either during ongoing motion perception (Muckli, Kohler, Kriegeskorte, & Singer, 2005; Yantis & Nakama, 1998), or during perceptual "filling in" in the brief interval directly after its sudden offset (Ekman et al., 2017).

Experiment 2b tested whether the perceptual shifts towards the efficient trajectories recorded in Experiments 1a-c can be disrupted with a short dynamic visual noise mask presented directly after action offset. Such masks reliably disrupt lower-level perceptual processes (Kinsbourne & Warrington, 1962; Breitmeyer & Ögmen 2006),

eliciting similar effects as transcranial magnetic stimulation of occipital cortices (Tapia & Beck, 2014). If masking successfully reduces predictive effects, then this would further confirm that the perceptual biases in Experiments 1a-c reflect changes to early visual stimulus representation, either “on-line” during action observation, or in the brief interval after its offset where expected trajectories are spontaneously “filled in” (Ekman et al., 2017).

The experiments in this chapter were published in Proceedings of the Royal Society: Biological Sciences, and form part of the supplementary experiments that support the main findings from the experiments in Chapter Three.

4.1 Using probe judgements and visual noise masking to reveal perceptual predictions

Human and non-human primates take the “intentional stance” when watching conspecifics (Dennett, 1987). They interpret the behaviour of others as goal directed (Baillargeon, Scott & Bian, 2016; Baker, Saxe & Tenenbaum, 2009; Csibra & Gergely, 2007; Gergely & Csibra, 2003) and expect them to achieve goals in the most efficient and rational way, minimising time and energy expenditure given the environmental constraints (Gergely & Csibra, 2003; Rochat, Serra, Fadiga & Gallese, 2008). This simple heuristic of action efficiency arises early in development and allows children to attribute intentionality to observed behaviours, even when carried out by inanimate objects (Gergely, Bekkering, & Király, 2002; Gergely, Nádasdy, Csibra, & Bíró, 1995; Liu & Spelke, 2017), forming a stepping-stone for more sophisticated abilities for reasoning about others (Gergely & Csibra, 2003; Wellman & Brandone, 2009).

The experiments in Chapter 3 reveal that the principle of efficient action guides the perceptual representation of others actions towards the most efficient trajectory towards goals, given the environmental constraints (Hudson, McDonough, Edwards, & Bach, 2018). They show that expectations of efficient action are, to some extent, perceptually represented, in the form of an ideal “reference” trajectory that a rational actor would take through a given environment, against which observed actions can be judged (Bach & Schenke, 2017; Hudson et al., 2018). Participants observed short video clips of a hand starting to reach for an object with a straight or arched trajectory. The actions were either efficient (reaching straight when the path was clear or arched over an obstacle) or inefficient (straight towards an obstacle or arched over empty space). Immediately after action offset, participants reported the hand’s last seen position on a touch screen. The results revealed perceptual biases towards efficient action expectations, such that straight reaches were reported to have reached higher when an obstacle was blocking its path, and arched reaches were reported lower when no obstacle was present. These distortions were elicited spontaneously but increased as predictions were made more explicit. Such effects are akin to the perceptual “filling in” of missing input by top-down information (Muckli et al., 2005; Yantis & Nakama, 1998) and the perceptual sharpening of motion perception during uncertainty (Bex, Edgar, & Smith, 1995; Ekman et al., 2017). Together, these results indicate that the teleological stance is at least partly perceptually represented, providing an ideal reference trajectory that informs the action that was indeed perceived.

In the study described above, perceptual judgments were recorded using touch screen localisations. These touch screen responses provide a direct measure of perceptual shift in each trial (Hubbard, 2005), but leave open at which processing

step they occur. Do they directly affect the perceptual representations of the observed actions, or do they emerge from later changes to the action's perceptual representations in working memory or in the sensorimotor maps that guide the motor responses to the relevant locations on the screen (Firestone & Scholl, 2015, Kerzel, 2005)? The experiments in this chapter address this question by replicating the main experiments with a probe comparison design that is not influenced by changes in memory or motor related processes (Kerzel, 2005), and with a visual noise masking paradigm that reliably interrupts perceptual processing (Breitmeyer & Öğmen, 2006). If the findings from the main experiments truly reflect changes to perceptual representations from expectations of action efficiency, then we would expect to replicate the predictive effects within the probe design and eliminate the effects with a dynamic visual noise mask.

4.2 Experiment 2a

The main experiments measured perceptual shifts by calculating the difference between the final location of the action and where participants judged this location to be, using a touch screen monitor (Hudson et al., 2018). This measure of response has been used in previous studies of predictive displacement (Hudson et al., 2017; Kerzel & Gegenfurtner, 2003; Motes, Hubbard, Courtney, & Rypma, 2008) and utilises the advances in technology that out-date previous Representational Momentum designs that used cursor positioning to capture localisation judgements (Hubbard & Bharucha, 1988; Joordens, Spalek, Razmy, & van Duijn, 2004). Both cursor positioning and touch screen responses provide a more direct measure of displacement than classic probe comparison designs used in original Representational Momentum paradigms (Freyd & Finke, 1984, 1985; Hubbard,

1997, 2005). They allow for the specific localisation for a given vanishing point in a single trial whereas probe designs would require several presentations of the same vanishing point with varying probe positions, even then only retrieving an estimate of this localisation.

Probe judgments, in contrast, have the advantage of better controlling for response time, minimising the delay incurred by localising the cursor or finger in the intended position, and, as they only require a single button press, minimise the spatial and motor components of the response (Hubbard, 2005). Replicating the original experiments with a Probe comparison design can therefore eliminate the influence of working-memory components, by controlling the retention interval between action offset and response, and motor related processes by employing the same response method (button press) in every trial that do not require access to visuospatial information. Since the original touch screen experiments have already identified the average position of predictive displacements, replications with a probe design are no longer limited by non-specific estimations for probe positions.

Experiment 2a therefore replicated the Report Object experiment from the original study (Hudson et al., 2018) with the well-established probe comparison task that is free from such memory or motoric influences, but reliably measures changes to the perceived motion in the predicted path (i.e., representational momentum, Freyd & Finke, 1984; Hudson et al., 2017; 2016ab, for reviews, see Hubbard, 2005; Kerzel, 2005). In each trial, participants compared the hand's last seen position with a probe stimulus presented directly after hand offset (250 ms. gap), which was displaced vertically either in the predicted direction (e.g. downwards for inefficient arched reaches) or in the opposite unpredicted direction, and horizontally leftwards or rightwards. They indicated, with the press of a button, whether the probe stimulus

position was identical or different from the hand's last seen position on the screen. Importantly, if predictions of efficient action affect the ongoing perceptual representation of the observed actions, akin to perceptual processing in non-biological perception (Muckli et al., 2005; Yantis & Nakama, 1998) or if they lead to spontaneous perceptual filling in of the predicted trajectories after the sudden offset (Ekman et al., 2017), then participants should be more likely to mistake probe displacements in the expected direction with the hand's last seen position, compared to displacements in the opposite, unpredicted direction. Because the probe stimuli appear directly after action offset and participants' responses do not need access to visuospatial representations, any such effects will therefore reflect either perceptual changes during ongoing action observation or directly after action offset.

4.2.1 Method

4.2.1.1 Participants

Thirty-nine participants took part in the experiment (mean age = 20.0 years, SD = 1.7, 28 females). All participants were right-handed, had normal/corrected vision, and were recruited from Plymouth University for course credit. The study received ethical approval from the University of Plymouth's ethics board, in accordance with those of the ESRC and the Declaration of Helsinki.

4.2.1.2 Apparatus

The experiment was presented on a HP EliteDisplay S230tm 23-inch widescreen (1920 X 1080) touch screen monitor. Verbal responses were recorded

with Microsoft LifeChat LX-3000 Headsets. All other components of the apparatus were the same as in the main experiments.

4.2.1.3 Stimuli

The stimulus set was identical to the main experiments. The only addition was the probe stimulus, a single red circle the same size (30 X 30 pixels) as the tip of the index finger of the action stimuli in Experiments 1a-c.

4.2.1.4 Procedure

The design of the experiment closely matched that of Experiments 1a-c (see Figure 4-1A). As before, participants completed two blocks of 80 randomised trials. Each trial began with the first static image of the action sequence, and continued to replicate the trial sequence of the Report Obstacle experiment (Experiment 1b) until the response stimulus. Thus, participants saw the action commence after they reported, verbally into the microphone, whether an obstacle was present in the scene. After the action disappeared, participants did not make a touch response. Instead, the probe stimulus was presented 250ms after hand offset (preventing masking effects, Breitmeyer & Öğmen, 2006). The probe stimulus was overlaid on top of the scene (without hand) and was positioned at either the same coordinates as the tip of the index finger, or at one of 12 different positions (see Figure 4-1B). These positions were derived from the average displacements induced by inefficient compared to efficient actions recorded in the Report Obstacle experiment (Experiment 1b, X = 24.3 pixels, Y= 19.2 pixels). Four of the different positions were calculated as the coordinates of the tip of the index finger, plus or minus 50% of these average displacement pixels, 100% of the average displacements, or 150%.

Panel B of Figure 4-1 depicts all 13 possible probe positions. Participants were required to press the spacebar if they judged the probe to be in a position different to the tip of the index finger and do nothing if they judged it to be in the same position.

Each participant received two practice blocks containing six trials each. In the first practice block, the final action frame remained on screen instead of the response stimulus, and the probe was overlaid on top of this frame. This made it clear to participants when the probe was in the same or different position as the tip of the index finger. The second practice block was the same as the experimental trials.

4.2.2 Results

Participants were excluded if the correlation between their probe judgements and the probe positions was more than 3SD away from the median r value (X axis: median .858, *SD* = .141; Y axis: median =.898, *SD* = .123, 2 participants excluded). Exclusion of these participants does not affect the results. Individual trials were excluded if response times were faster than 200ms or slower than 3000ms (.04% of trials).

Analysis was conducted on the proportion of “different” responses, averaged across the three probe positions in each of the four directions. Difference scores were calculated along the X and Y axis separately to measure the size of the perceptual shift. For the X axis, responses for rightward probes were subtracted from responses for leftward probes. Therefore, positive difference scores denote the proportion of rightward probes judged as “same” and negative difference scores denote the proportion of leftward probes judged as “same”. For the Y axis, responses for upward probes were subtracted from responses for downward probes. Therefore, positive difference scores denote the proportion of upward probes judged

as “same” and negative difference scores denote the proportion of downward probes judged as “same”. These difference scores were entered into two separate 2 X 2 ANOVAs with Trajectory (arched, straight) and Efficiency (efficient, inefficient) as within-subjects factors.

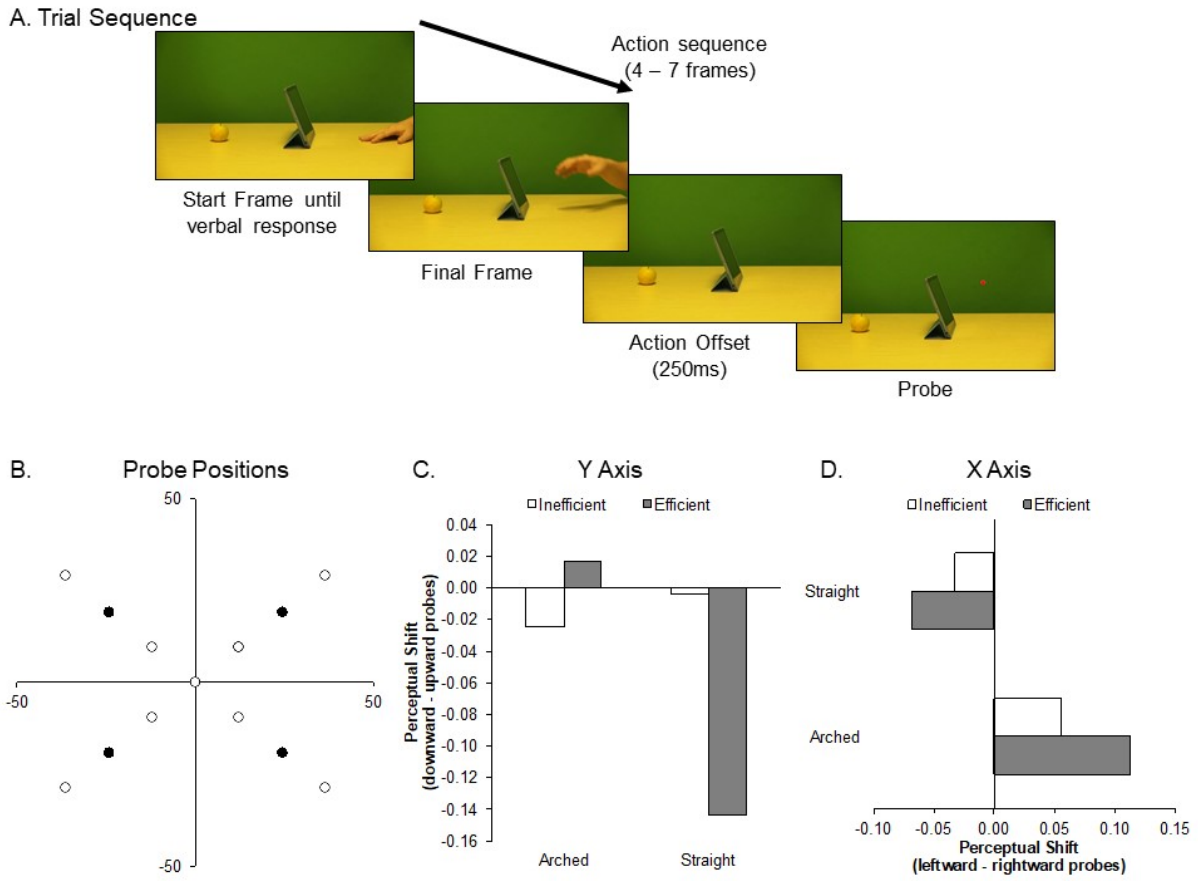


Figure 4-1. Experiment 2a. Trial sequence, Probe positions and Results.

An example of the trial sequence is depicted in Panel A. Panel B depicts all probe positions relative to the final position of the hand, where 0,0 depicts a probe in the same position as the hand. Filled circles depict average displacement pixels as recorded in the Report Obstacle Experiment and empty circles depict plus or minus 50%. The results for the Y axis are depicted in Panel C and the results for the X axis are depicted in Panel D.

4.2.2.1 Y Axis

Overall, there was a main effect of Trajectory (F(1,36)= 13.44, p=.001, η_p^2=.272), where the likelihood to accept upward compared to downward probes as "same" was greater for arched reaches (-.004) than for straight reaches (-.07, t(36)=3.67, p=.001, d=0.60), consistent with a extrapolation of the prior motion along its path. A main effect of Efficiency (F(1,36)= 5.66, p=.023, η_p^2=.136) indicated that the likelihood to accept upward compared to downward probes as "same" was greater for inefficient reaches (-.01) than for efficient reaches (-.06, t(36)=2.38, p=.023, d=0.39). Most importantly, the analysis revealed the predicted interaction between Trajectory and Efficiency (F(1,36)= 11.39, p=.002, η_p^2=.240). Participants were more likely to accept downwards compared to upwards probes as "same" for inefficient arched reaches than for efficient arched reaches, and, conversely, were more likely to accept upwards probes for inefficient straight reaches than for efficient straight reaches. These results therefore fully replicate the perceptual shifts towards the efficient trajectories in Experiments 1a-c with a psychophysical judgment task without working memory or motor components.

4.2.2.2 X Axis

As in Experiments 1a-c, we did not have specific predictions for the X Axis. The reported effects should therefore be considered exploratory and interpreted with caution. Overall, there was a main effect of Trajectory (F(1,36)= 47.01, p<.001, η_p^2=.566). The likelihood to accept rightward compared to leftward probes as "same" was greater for arched reaches (.08) than for straight reaches (-.05, t(36)=6.86, p<.001, d=1.23), most likely reflecting a greater expectation of forward momentum

(leftward direction) for straight reaches compared to arched reaches (Representational Momentum; Hubbard, 2005). Interestingly, the analysis revealed an interaction between Trajectory and Efficiency ($F(1,36)= 7.49$, $p=.010$, $\eta_p^2=.172$), showing that the likelihood to accept rightwards compared to leftward probes as "same" was greater for efficient arched reaches than for inefficient arched reaches, and greater for inefficient straight reaches than for efficient straight reaches. While unpredicted, this finding is fully in line with the expected deviation towards the predicted "efficient" trajectory. Because straight reaches exert more forward displacements than arched reaches (see above), this forward displacement also takes place – albeit to a smaller extent – when participants see an arched reach but predict a straight reach, or conversely, is reduced when participants see a straight reach but predict an arched one. As noted, this effect was not predicted and not observed with the touch screen responses. It should therefore be interpreted with caution before being replicated.

4.2.3 Discussion

The results of Experiment 2a confirm that perceptual distortions of observed actions towards an ideal reference trajectory can be measured with probe stimuli, with responses that do not rely on perceptual working memory representations or visuospatial motor maps (Kerzel, 2005). Participants simply reported – with a press of a button – whether the index finger's seen disappearance was identical to a probe stimulus presented directly after action offset, which could be displaced either subtly upwards or downwards from the real disappearance point. Mirroring the results of the main experiments (Hudson, et al., 2018), participants were more likely to mis-identify probes displaced towards the predicted "ideal" trajectory with the actually

perceived disappearance point. They more readily accepted downwards probes as the same as the last seen position of inefficient arched reaches, and upwards probes for inefficient straight reaches.

This replication rules out that the effects emerge from perceptual changes to the action's representation in later working memory or motor control stages, and instead reveal a contribution to immediate perceptual processing, either during ongoing motion perception (Muckli et al., 2005; Yantis & Nakama, 1998), or in the brief interval directly after its sudden offset, when the visual system spontaneously fills in the further expected trajectory (Ekman et al., 2017). These effects can be linked either to top-down changes that sharpen the considerable perceptual uncertainty during motion perception (motion blurring & sharpening, Hammett, 1997), and/or to changes in short term iconic memory that are assumed to rely on early visual representation. These effects are comparable to such phenomena as integration of stimulus features, change blindness and the experience of stable percepts across saccades (Becker, Pashler & Anstis, 2000; Jonides, Irwin & Yantis, 1982, see Öğmen & Herzog, 2016 for a recent review).

4.3 Experiment 2b

Experiment 2a provides strong support for the top-down influence of action efficiency expectations on the perceptual representation of others' actions. Replicating the findings from the main experiments (Hudson et al., 2018) with a probe comparison task, Experiment 2a discounts that effects could have emerged from perceptual changes to the action's representation in later working memory or motor control stages, and instead reveal a contribution to immediate perceptual processing, either

during ongoing motion perception (Muckli, et al., 2005; Yantis & Nakama, 1998), or in the brief interval directly after its sudden offset, when the visual system spontaneously fills in the further expected trajectory (Ekman et al., 2017). If predictions of efficient kinematics do indeed act on early perceptual representations, then methods that reliably disrupt these lower-level perceptual processes should equally disrupt the influence of action efficiency predictions on social perception.

It is a well-established phenomenon that visual masks, presented directly after stimulus offset (backward masking), reliably disrupt lower-level processing in perceptual regions (Breitmeyer & Ögmen 2006; Kinsbourne & Warrington, 1962). In backward masking paradigms, presenting a second visual stimulus (the mask) immediately after an initial visual stimulus (the target) impairs perception of this first target stimulus (Lamme 2000; Lamme & Roelfsema, 2000; Lamme, Zipser & Spekreijse, 2002). These studies showed that visual masking leaves early feedforward signals from primary visual cortex (V1) to higher areas that encode the visual stimulus features unaffected, while specifically disrupting the later re-entrant feedback signals from higher visual areas to V1, responsible for establishing a stable stimulus representation (e.g. “interruption masking”, Kolers, 1968; for more recent studies, see Boehler, Schoenfeld, Heinze, & Hopf, 2008; Fahrenfort, Scholte, & Lamme, 2007). They are further supported by studies that apply transcranial magnetic stimulation (TMS) to early visual cortices at the respective points reflecting feedforward and feedback processes, and induce comparable masking-like effects (for a review, see Tapia & Beck, 2014).

Thorough investigations have demonstrated that the disruption of top-down influences on visual processing are greatest when masks are presented immediately after target offset (SOAs from 30 to 100 ms) and when masks are more visually

intrusive (Breitmeyer & Ögmen 2006). Visually intrusive masks consist of high contrast random noise patterns (Kinsbourne and Warrington 1962), such as random black and white dot patterns that are dynamically presented (e.g. rapid sequential presentations of different random noise arrays, Quinn & McConnell, 1996). Importantly for the current experiment, such dynamic visual noise masks induce apparent motion illusions between the successive presentations of contrasting dots in different spatial locations, which specifically interfere with visual processing during motion perception (Breitmeyer & Ögmen 2006; MacKay, 1965).

Visual masking paradigms therefore induce specific interferences to top-down visual influences on perception, while retaining bottom-up sensory information from the stimulus (Fahrenfort et al., 2007, 2008), with dynamic visual noise masks creating unique disruptions to the visual processing of moving targets (Breitmeyer & Ögmen 2006; MacKay, 1965). They can interfere with visual processing during actual perception (e.g., backwards masking, Lamme 2000; Lamme & Roelfsema, 2000; Lamme, Zipser & Spekreijse, 2002) or during visual imagery, where masking interferes with the “painting” of top-down information into perceptual structures (Andrade, Kemps, Werniers, May, & Szmalec, 2002; Borst, Ganis, Thompson & Kosslyn, 2012; McConnell & Quinn, 2000; Quinn and McConnell, 1996, 1999).

To test whether such top-down interactions with early visual processes are responsible for the biases towards efficient actions, we again replicated the Report Obstacle experiment (Hudson et al., 2018) but inserted, in half of the trials, a short (560 ms.) rapidly changing visual noise pattern immediately after the action offset. Because such dynamic visual noise causes apparent motion (MacKay, 1965), it should interfere with motion based predictions that contribute either to the conscious perception of the seen action, or to the perceptual “filling in” of the suddenly missing

information directly after action offset. If the perceptual biases emerge from such changes to early visual perceptual representations, then these biases should be only (or more strongly) observed in the no-mask compared to the masked trials.

4.3.1 Method

4.3.1.1 Participants

Twenty-eight participants took part in the experiment (mean age = 19.6 years, SD = 1.1, 26 females). All participants were right-handed, had normal/corrected vision, and were recruited from Plymouth University for course credit. The study received ethical approval from the University of Plymouth's ethics board, in accordance with those of the ESRC and the Declaration of Helsinki.

4.3.1.2 Apparatus

The experiment was presented on a HP EliteDisplay S230tm 23-inch widescreen (1920 X 1080) touch screen monitor. Verbal responses were recorded with Microsoft LifeChat LX-3000 Headsets. All other components of the apparatus were the same as in the main experiments.

4.3.1.3 Stimuli

The stimulus set was identical to the main experiments. The additional mask stimuli were created in R. The mask covered an area of 200 X 200 pixels and contained 50 black and 50 white squares of equal size (12 X 12 pixels) on a transparent background. Twenty different mask images were created, each containing a randomised arrangement of the squares.

4.3.1.4 Procedure

The design of the experiment closely matched the Report Obstacle version of the main experiments. Participants completed two blocks of 80 randomised trials. Half of the trials were an exact replication of the Report Obstacle experiment (no-mask condition), and half the trials had the addition of the mask (mask condition), randomly interspersed. Participants again reported whether an obstacle was present in the scene or not, by speaking “Yes” or “No” into the microphone. The action sequence then started and disappeared before completion. In no-mask trials, participants simply indicated on the response stimulus – the scene with the hand removed – the index finger’s last seen location. For masked trials, the mask was overlaid on top of the response stimulus 560 ms immediately after action offset, on which participants reported – with a touch response – the hand’s last seen position. The centre of the mask was positioned at the disappearance point of the tip of the index finger, plus or minus 20 pixels in the X and Y direction, to ensure that participants could not simply use the task to aid their judgment. As soon as the hand disappeared, a sequence of seven randomised mask images was presented at the same rate as the prior action sequence (80 ms. per frame), creating a mask which was on screen for 560 ms. Once the mask ended, the response stimulus remained on screen until the touch response was recorded. Any touch responses recorded while the mask remained on screen ended the trial. An example trial sequence for masked trials can be seen in Figure 4-2A.

4.3.2 Results

Exclusion criteria were identical to the main experiments. No participants were excluded on the basis of the distance between the real and selected screen coordinate (mean = 36.3px, *SD* = 21.9), but one was excluded based on the correlation between the real and selected positions on the X (median r = .944, *SD* = .039) or Y axis (median r = .888, *SD* = .038). A total of 3.2% of trials were excluded due to incorrect response procedure and 2.8% of trials were excluded if initiation or execution times were less than 200ms or more than 3SD above the sample mean (Initiation: mean =350.5ms, *SD*=158.7; Execution: mean =527.8ms, *SD*=161.8). In 2.9% of trials, a response was made while the mask remained on screen. These trials were included in the analysis but their exclusion/inclusion does not affect the results.

Data was analysed in the same way as the main experiments. Difference values (reported minus actual disappearance points) were entered into a 2X2X2 repeated-measures ANOVA for the X and Y coordinates separately, with Trajectory (arched, straight), Efficiency (efficient, inefficient), and Condition (mask, no-mask) as within-subjects factors.

4.3.2.1 Y Axis

Overall, there was a main effect of Trajectory ($F(1,26)= 80.08$, $p<.001$, $\eta_p^2=.755$) where arched reaches (-1.8px) were displaced higher than straight reaches (-9.7px, $t(26)=8.74$, $p<.001$, $d=1.68$). Importantly, the analysis replicated the interaction of Efficiency and Trajectory ($F(1,26)= 22.74$, $p<.001$, $\eta_p^2=.467$). As before, inefficient arched trajectories (-4.5px) were displaced below efficient arched trajectories (0.7px,

$t(26)$=-3.71, p=.001, d=.71), and inefficient straight trajectories (-7.3px) were displaced above efficient straight trajectories (-12.1px, , $t(26)$=4.25, p<.001, d=.82). More importantly, the analysis revealed a three-way interaction between Efficiency, Trajectory and Mask ($F(1,26)$=8.89, p=.006, η_p^2=.255). As predicted, while the displacements towards the more effective trajectory in the no-mask trials closely resembled the main experiment (14.4 vs. 14.6 pixels, respectively), they were substantially reduced in the masked trials (5.2 pixels, see Figure 4-2D).

4.3.2.2 X Axis

Overall, there was a main effect of Trajectory ($F(1,26)$= 148.85, p<.001, η_p^2=.852). As in the main experiments, arched trajectories (-25.3px) were displaced more leftward than straight trajectories (-4.5px, $t(26)$=-11.98, p<.001, d=2.3). An interaction between Trajectory and Mask ($F(1,26)$= 41.41, p<.001, η_p^2=.614) showed that the Trajectory effect was larger in the no-mask trials than in the masked trials ($t(26)$=6.44, p<.001, d=1.2). There was also a three-way interaction between Efficiency, Trajectory and Mask condition ($F(1,26)$=4.89, p=.036, η_p^2=.158) revealing that the Trajectory X Mask condition effect was larger for Efficient actions than for Inefficient actions. While this effect reveals a similar mask effect as for the Y Axis, it should be treated with caution as it was not predicted, no similar interaction of Efficiency and Trajectory was found for any of the main experiments, and it was one of many possible (unpredicted) effects in the ANOVA, and would therefore be subject to adjustments for multiple comparisons (Cramer et al., 2016).

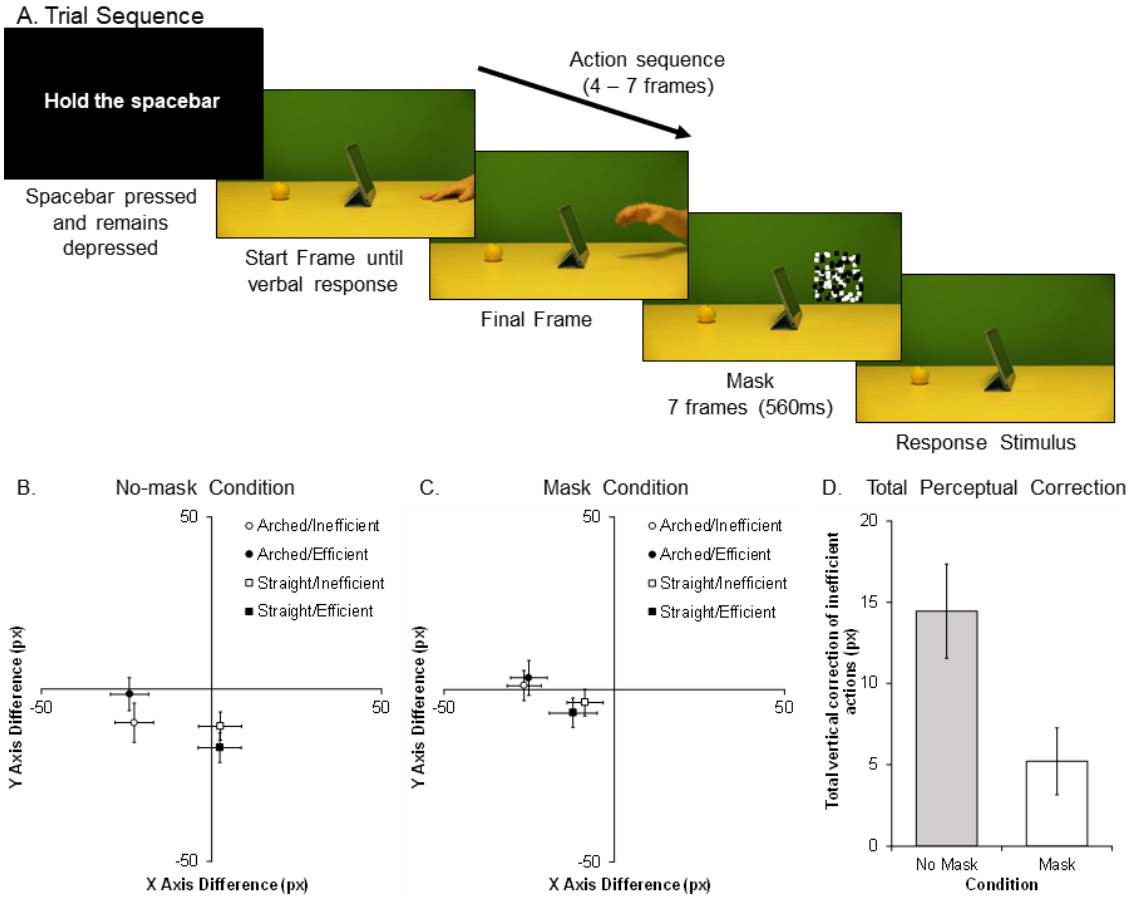


Figure 4-2. Experiment 2b. Trial sequence and Results.

An example of the trial sequence for the Mask condition is depicted in Panel A. The results for the no-mask condition are depicted in Panel B and the results for the Mask condition are depicted in Panel C. Panel D depicts a comparison of the size of the Y axis interaction in pixels, equivalent to the total amount by which inefficient actions were corrected towards a more efficient trajectory. Error bars represent 95% confidence intervals.

4.3.3 Discussion

Experiment 2b replicated the finding that perceptual judgments of observed actions are biased towards efficient trajectories. Crucially, it showed that a brief dynamic visual noise mask inserted directly after action offset successfully disrupted the resulting effects on perceptual judgments, substantially reducing the bias towards efficient actions. Dynamic visual noise masks as used here specifically interfere with

the re-entrant top-down interactions with early perceptual regions (Boehler et al., 2008; Fahrenfort, Scholte, & Lamme, 2007) that are crucial for visual awareness of a stimulus (Lamme & Roelfsema, 2000; Lamme et al., 2002), especially during motion perception (Breitmeyer & Ögmen 2006; MacKay, 1965), or the creation of a detailed mental image during visual imagery that is akin to actual perception and which can be accessed for further processing (Andrade et al., 2002; Borst et al., 2012; McConnell & Quinn, 2000). The masking effects therefore further confirm that the perceptual bias in the main experiments either reflect on-line changes to the action's perceptual representation during observation, or spontaneous "filling in" of the suddenly missing input briefly after its offset, creating an impression of an action displaced towards the anticipated ideal reference trajectory.

4.4 Conclusion

Experiments 2a and 2b showed that the biases in perceptual judgements found in the main experiments (Hudson et al., 2018) could also be observed in a probe judgment task without working memory or motor components already at 250 ms. after action offset. Moreover, they showed that the perceptual biases were effectively disrupted by dynamic visual noise masks which interfere with the re-current (top-down) feedback to early visual cortex (Fahrenfort, Scholte, & Lamme, 2007), preventing its use in awareness (i.e. backwards masking, Breitmeyer & Ögmen 2006; Tapia & Beck, 2014) or visual imagery (e.g., Andrade et al., 2002; Borst et al., 2012). The observed biases in perceptual judgments are therefore unlikely to stem from unspecific perceptual changes in memory or motor control (e.g. Firestone & Scholl, 2016; see Ianì, Mazzoni & Bucciarelli, 2018, for an example for perceptual changes in action memory), instead supporting a role in ongoing motion

perception (e.g. Muckli et al., 2005) or occurring directly after its offset, when the predicted future trajectory is visually “filled in” (e.g. Ekman et al., 2017). Together with the findings from Chapter Three, these results reveal that social perception is driven by similar predictive mechanisms as non-social perception, such as the assumption that light comes from above (Adams et al., 2004) or that gravity pulls objects downwards (Hubbard & Bharucha, 1988). Specifically, the principle of efficient action provides a perceptual reference signal that guides the perceptual representation of inefficient actions towards a more rational path. The resulting biases in perceptual judgments reveal concrete expectations of the specific trajectory that the action will take through the scene. Action efficiency is therefore constantly projected onto other’s behaviour via predictions that can be derived from contextual cues prior to the motor behaviour.

4.5 References

Andrade, J., Kemps, E., Werniers, Y., May, J., & Szmalec, A. (2002). Insensitivity of visual short-term memory to irrelevant visual information. *The Quarterly Journal of Experimental Psychology: Section A, 55,* 753-774. doi:10.1080/02724980143000541

Becker, M. W., Pashler, H., & Anstis, S. M. (2000). The role of iconic memory in change-detection tasks. *Perception, 29,* 273-286. doi:10.1068/p3035

Boehler, C. N., Schoenfeld, M. A., Heinze, H. J., & Hopf, J. M. (2008). Rapid recurrent processing gates awareness in primary visual cortex. *Proceedings of the National Academy of Sciences, 105,* 8742-8747. doi:10.1073/pnas.0801999105

Borst, G., Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2012). Representations in mental imagery and working memory: Evidence from different types of visual masks. *Memory & Cognition, 40,* 204-217. doi:10.3758/s13421-011-0143-7

Breitmeyer, B., & Öğmen, H. (2006). Visual masking: Time slices through conscious and unconscious vision (No. 41). Oxford University Press.

Cramer, A. O., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P., ... & Wagenmakers, E. J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review, 23,* 640-647. doi:10.3758/s13423-015-0913-5

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications, 8,* 15276. doi:10.1038/ncomms15276

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 126-132. doi:10.1037/0278-7393.10.1.126

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience, 19,* 1488-1497. doi:10.1162/jocn.2007.19.9.1488

Hammett, S. T. (1997). Motion blur and motion sharpening in the human visual system. *Vision Research, 37,* 2505-2510. doi:10.1016/S0042-6989(97)00059-X

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12,* 822-851. doi:10.3758/BF03196775

Hudson, M., Bach, P., & Nicholson, T. (2017). You said you would! The predictability of other's behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance, 44*, 320-335. doi:10.1037/xhp0000451

Hudson, M., Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition, 146,* 245-250. doi:10.1016/j.cognition.2015.09.021

Hudson, M., Nicholson, T., Simpson, W. A., Ellis, R., & Bach, P. (2016). One step ahead: The perceived kinematics of others' actions are biased toward expected goals. *Journal of Experimental Psychology: General, 145,* 1-7. doi:10.1037/xge0000126

Jonides, J., Irwin, D. E., & Yantis, S. (1982). Integrating visual information from successive fixations. *Science, 215,* 192-194. doi:10.1126/science.7053571

Kerzel, D. (2005). Representational momentum beyond internalized physics: Embodied mechanisms of anticipation cause errors in visual short-term memory. *Current Directions in Psychological Science, 14,* 180-184. doi:10.1111/j.0963-7214.2005.00360.x

Kinsbourne, M., & Warrington, E. K. (1962). The effect of an after-coming random pattern on the perception of brief visual stimuli. *The Quarterly Journal of Experimental Psychology, 14,* 223-234. doi:10.1080/17470216208416540

Lamme V. A. F. (2000). Neural mechanisms of visual awareness: A linking proposition. *Brain and Mind, 1,* 385–406. doi:10.1023/A:1011569019782

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23,* 571-579. doi:10.1016/S0166-2236(00)01657-X

Lamme, V. A., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience, 14,* 1044-1053. doi:10.1162/089892902320474490

MacKay, D. M. (1965). Visual noise as a tool of research. *The Journal of General Psychology, 72,* 181-197. doi:10.1080/00221309.1965.9710688

McConnell, J., & Quinn, J. G. (2000). Interference in visual working memory. *The Quarterly Journal of Experimental Psychology: Section A, 53,* 53-67. doi:10.1080/713755873

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology, 3,* e265. doi:10.1371/journal.pbio.0030265

Öğmen, H., & Herzog, M. H. (2016). A new conceptualization of human visual sensory-memory. *Frontiers in Psychology, 7,* 830. doi:10.3389/fpsyg.2016.00830

Quinn, J. G., & McConnell, J. (1996). Irrelevant pictures in visual working memory. *The Quarterly Journal of Experimental Psychology Section A, 49,* 200-215. doi:10.1080/713755613

Quinn, J. G., & McConnell, J. (1999). Manipulation of interference in the passive visual store. *European Journal of Cognitive Psychology, 11,* 373-389. doi:10.1080/713752322

Tapia, E., & Beck, D. M. (2014). Probing feedforward and feedback contributions to awareness with visual masking and transcranial magnetic stimulation. *Frontiers in Psychology, 5,* 1173. doi:10.3389/fpsyg.2014.01173

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience, 1,* 508-512. doi:10.1038/2226

5 Chapter Five - Cues to intention

The experiments in Chapter Three and Four have demonstrated how the principle of efficient action allows observers to spontaneously predict ideal reference trajectories that a rational and intentional actor will follow, given that they are fully aware of all relevant environmental constraints. These predictions are at least partially perceptually represented and specifically concern the expected action kinematics, biasing perception towards these expectations. This mechanism of predictive social perception is in line with recent predictive processing theories (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007ab) and is comparable to the mechanisms that underlie non-social perception (Clark, 2013; Friston, 2010). What remains unresolved is how these efficiency predictions emerge and upon what stimulus features they depend. If assumptions of efficient action are only assigned to intentional agents, then formation of such predictions should only emerge when cues to the intentionality of an actor are present.

The experiments in Chapter Five investigate the cues that trigger attributions of intentionality to others and whether the expectation of efficient action is tied to such intentionality attribution. The findings are important not only for supporting the proposal that these predictions emerge from the attribution of intention to others, but also for emphasising how revisions to these attributions, when observing inefficient actions, can signal that a behaviour may not have been intentional after all, or that the actor is not aware of all relevant environmental constraints. Such findings will therefore reinforce the proposed link between teleological perception of others' behaviour and more sophisticated theory of mind and mentalising processes, contributing to action understanding.

The experiments in this chapter were published in Scientific Reports and are presented in their published format (green copy).

5.1 Cues to intention bias action perception toward the most efficient trajectory.

Katrina L. McDonough, Matthew Hudson, Patric Bach

Abstract

Humans interpret others' behaviour as intentional and expect them to take the most energy-efficient path to achieve their goals. Recent studies show that these expectations of efficient action take the form of a prediction of an ideal "reference" trajectory, against which observed actions are evaluated, distorting their perceptual representation towards this expected path. Here we tested whether these predictions depend upon the implied intentionality of the stimulus. Participants saw videos of an actor reaching either efficiently (straight towards an object or arched over an obstacle) or inefficiently (straight towards an obstacle or arched over an empty space). The hand disappeared mid-trajectory and participants reported the last seen position on a touch-screen. As in prior research, judgments of inefficient actions were biased toward efficiency expectations (straight trajectories upwards to avoid obstacles, arched trajectories downward towards goals). In two further experimental groups, intentionality cues were removed by replacing the hand with a non-agentive ball (group 2), and by removing the action's biological motion profile (group 3). Removing these cues substantially reduced perceptual biases. Our results therefore confirm that the perception of others' actions is guided by expectations of efficient actions, which are triggered by the perception of semantic and motion cues to intentionality.

5.2 Experiments 3a and 3b

Humans see others' behaviour as purposeful and goal directed (Baillargeon, Scott, & Bian, 2016; Baker, Saxe, & Tenenbaum, 2009; Csibra & Gergely, 2007; Gergely & Csibra, 2003). A key signature of this "intentional stance" (Dennett, 1987) is the assumption that other people generally act rationally: they take the most energy-efficient path to achieve their goal, and expend additional energy only when an obstacle has to be overcome (Csibra & Gergely, 2007; Hunnius & Bekkering, 2014). This simple heuristic of action efficiency arises early in development and allows children to attribute intentionality to observed behaviours, even when carried out by inanimate objects (Gergely, Bekkering, & Király, 2002; Gergely, Nádasdy, Csibra, & Bíró, 1995; Liu & Spelke, 2017). Human infants (and some non-human primates) show surprise, for example, when actors that are believed to be intentional violate these assumptions, such as when they do not adjust their reach despite an obstacle or exert additional unnecessary energy to reach their goal (Gergely & Csibra, 2003; Rochat, Serra, Fadiga, & Gallese, 2008). Once established, this simple heuristic may form a stepping-stone for more sophisticated abilities for reasoning about others (Gergely & Csibra, 2003; Wellman & Brandone, 2009). For example, observing an inefficient action (e.g. reaching directly towards an object despite an obstacle in the way) can help people realize that others act according to beliefs and not objective reality (i.e. they may not have sight of the obstacle), forming the basis of a prototypical theory of mind.

We have argued that expectations of efficient action are, to some extent, perceptually represented, in the form of an ideal "reference" trajectory that a rational actor would take through a given environment, against which observed actions can

be judged (Bach & Schenke, 2017; Hudson, McDonough, Edwards, & Bach, 2018). This proposal emerges from recent predictive processing models of social perception (Bach & Schenke, 2017; Csibra, 2008; Kilner, Friston, & Frith, 2007a, 2007b; Zaki, 2013) which argue that perception of others' actions – like perception in general – is hypothesis-driven. Any assumption about the external world (and the people within it) is translated into the perceptual input that would result from such a state. These expectations of future input can guide perception and be tested against actual stimulation (Clark, 2013; Friston & Kiebel, 2009; Hohwy, 2013). In non-social perception, such expectations explain several visual illusions (e.g., dress illusion, Schlaffke et al., 2015), the switch between different bi-stable percepts (Kondo, Farkas, Denham, Asai, & Winkler, 2017), or why the same objects can appear convex or concave depending on prior assumptions about light sources (Adams, Graf, & Ernst, 2004). In social perception, simply attributing a goal to another person could similarly elicit associated predictions about how this individual would realise such a goal, specifying which action they may soon carry out (Bach, Bayliss, & Tipper, 2011; Bach, Knoblich, Gunter, Friederici, & Prinz, 2005; for theoretical arguments, see Bach, Nicholson, & Hudson, 2014; Bach, Nicholson, & Hudson, 2015; Csibra, 2008; Kilner et al., 2007ab). The principle of efficient action can make a direct contribution here, specifying the ideal "reference" trajectory that achieves the actor's goals with minimum energy expenditure, given the current environmental constraints, such as potential obstacles in the way (Bach et al., 2014; Bach & Schenke, 2017).

In a recent series of studies, we attempted to reveal these expectations of efficient action (Hudson et al., 2018). These studies relied on the well-established phenomenon that the uncertainty during motion perception is perceptually sharpened

using top-down information (Bex, Edgar, & Smith, 1995; Hammett, 1997), filling in missing information (Muckli, Kohler, Kriegeskorte, & Singer, 2005; Shiffrar & Freyd, 1993; Yantis & Nakama, 1998) in a predictive manner (Avenanti, Annella, Candidi, Urgesi, & Aglioti, 2012; Ekman, Kok, & de Lange, 2017). The resulting perceptual biases can be reliably measured by suddenly removing the moving object from view, and asking participants to report its disappearance point, either on a touch screen (Pozzo, Papaxanthis, Petit, Schweighofer, & Stucchi, 2006; Saunier, Papaxanthis, Vargas, & Pozzo, 2008, see also, Hudson, Bach, & Nicholson, 2017) or by comparing it to probe stimuli presented shortly after (Freyd & Finke, 1984; Hubbard, 2005, 2015, see also, Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016). In such a paradigm, people generally over-estimate the movement they have seen, reporting the moving stimulus to have disappeared further along its trajectory than it really did (i.e. the Representational Momentum effect, Freyd & Finke, 1984; Hubbard, 2005). These displacements have been shown not only reflect a sample extrapolation of motion based on the previously seen trajectory (Kessler, Gordon, Cessford, & Lages, 2010), but also prior knowledge about it's causes, such as how the motion would be affected by one's own actions (Jordan & Hunsinger, 2008), by physical forces such as friction or gravity (Hubbard, 1995), or the most likely behaviours of the other person (Hudson, Burnett, & Jellema, 2012; Hudson & Jellema, 2011; Hudson, Liu, & Jellema, 2009; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016).

In the case of observed actions, the perceptual biases reflect the predictions derived from the assumption of efficient action (Hudson et al., 2018). In a recent series of experiments, participants observed a hand reaching for an object with a straight or arched trajectory. The actions were either efficient (reaching straight when the path

was clear or arched over an obstacle) or inefficient (straight towards an obstacle or arched over empty space). The movement disappeared at some point on its course and participants reported the hand's last seen position on a touch screen, or by comparing it to probe stimuli presented immediately after. Both measures revealed that perceptual judgements were reliably biased by expectations of efficient action. Straight reaches were reported to have reached higher when an obstacle was blocking its path, in line with the expectation that the hand would soon lift to avoid it. Conversely, high arched reaches were reported lower when no obstacle was present, and corrected towards the straighter, more energy-efficient trajectory. These biases were present automatically, but increased when participants explicitly predicted – prior to action onset – the most efficient trajectory through the scene, or when attention was drawn to the environmental constraints. Moreover, they could be disrupted by dynamic visual noise masks presented directly after stimulus offset, suggesting that the biases emerge during ongoing perception or directly after the sudden offset, when the visual system "fills in" the expected future path (Fahrenfort, Scholte, & Lamme, 2007; Lamme, Zipser, & Spekreijse, 2002).

Together, these results indicate that the teleological stance is at least partly perceptually represented, providing an ideal reference trajectory that informs the action that was indeed perceived. Here, we test on what stimulus features these predictions of efficient action depend. In children, as well as in adults, intention attribution – and the resulting surprise when seeing an inefficient action – depends on the presence of cues to intention (Johnson, 2000, 2003; Sartori, Becchio, & Castiello, 2011), such as seeing an agentive stimulus (such as a hand relative to a ball; Falck-Ytter, Gredebäck, & von Hofsten, 2006), or observing movements with biological motion trajectories (Baron-Cohen, 1997; Leslie, 1994; Morewedge,

Preston, & Wegner, 2007; Rakison & Poulin-Dubois, 2001). If such cues indeed trigger attributions of intentionality to others, and the expectation of efficient action is tied to such intentionality attribution, then they should also determine to what extent perceptual biases towards efficient actions are observed.

In the first experimental group, we replicated the original experiment by Hudson and colleagues (Hudson et al., 2018), in which participants saw efficient and inefficient reach trajectories (arched/straight over an obstacle vs. empty space) and indicated the hand's last seen location before it suddenly disappeared on a touch screen. In two further experimental groups, we progressively removed intentional cues. First, as in prior research on infant intention attribution, we replaced the hand with a non-agentive stimulus – a ball (Falck-Ytter et al., 2006) –, which however followed the same characteristic biological motion trajectories and profiles as the hands in the first experimental group, showing the classical bell-shaped velocity profile of reaches towards objects (Beggs & Howarth, 1972). Second, humans are sensitive to motion cues that distinguish the intentional biological agents from inanimate objects, such as self-propulsion and change of direction (Baron-Cohen, 1997; Leslie, 1994), or a trajectory and speed of movement that is similar to one's own movement (Morewedge et al., 2007; Rakison & Poulin-Dubois, 2001). In a third group, participants therefore saw the same ball, but it did not now follow a biological motion profile, removing all kinematic cues to intention. If biases toward efficient action emerge from cues that signal intentionality, then they should be substantially reduced in group 2, and further reduced in group 3, as cues to intentionality are removed.

5.2.1 Method

5.2.1.1 Participants

Eighty-two participants took part in the experiment: twenty-nine participants in group 1 (hand stimuli, mean age = 21 years, SD = 4.7, 25 females), twenty-seven in group 2 (balls with biological motion, mean age = 20 years, SD = 4.1, 21 females), and twenty-six in group 3 (balls with non-biological motion, mean age = 21 years, SD = 4.2, 20 females). Nine additional participants (two from group 1, three from group 2, four from group 3) were excluded based on previously established exclusion criteria (see Results). All participants in all groups were tested in the same three-week period. Both the gender mix and age distribution did not differ between groups (ps>.39). All participants were right-handed, had normal/corrected-to-normal vision, gave informed consent, and were recruited from the University of Plymouth or the wider community for course credit or payment. The study was approved by the University of Plymouth's ethics board, in line with the ESRC and the Declaration of Helsinki. A power analysis revealed that a sample size of 26 provides .80 power to detect two-sided within-subjects effects in each of the group with Cohen's *d* = .57. Our prior study investigating the same effect (Hudson, et al., 2018, "report obstacle" condition) and pilot data revealed consistently larger effect sizes, *d*= .76 to *d* = 1.29. For the between-subjects effects, a power analysis revealed that a sample size of 26 per group provides .80 power to detect effects in either direction with Cohen's *d* = .79. This should provide enough power to detect reductions of the original effect to about 40% of the original size (assuming that standard deviations remain the same).

5.2.1.2 Apparatus

Presentation (NeuroBS) software was used to present the experiment via a HP EliteDisplay S230tm 23-inch widescreen (1920 x 1080) Touch Monitor. Verbal responses were recorded with Presentation's sound threshold logic via a Logitech PC120 combined microphone and headphone set.

5.2.1.3 Stimuli

Example stimuli can be seen in Figure 5-1A-B. To derive a set of stimuli of efficient actions, videos were filmed of an arm at rest to the right of the screen, which then reached for one of four objects (an apple, a packet of crisps, a glue stick, or a stapler) on the left of the screen. The reaches were either directed straight for the target object (Straight/Efficient), or arched over one of three obstacles (an iPad, lamp or pencil holder; Arched/Efficient). Each video clip was then converted into individual frames, and the first 22 from frame 1 (initial rest position) to 22 (mid-way through the action) were used as stimuli. For each efficient action, an inefficient action sequence was created by digitally removing the obstacles from the Arched/Efficient videos (Arched/Inefficient), or by inserting the obstructing objects into the Straight/Efficient videos, (Straight/Inefficient). The inefficient actions were therefore identical to the efficient actions in terms of movement kinematics, and differed only by the presence/absence of the obstacle. Finally, response stimuli were created by taking one frame from each action sequence and digitally removing the actor's arm from the scene, so that only the objects and background remained. Presenting this frame immediately after the action sequence gave the impression of the hand disappearing

from the scene, and participants indicated the last seen location of the tip of the index finger on this frame with a touch response on screen.

For experimental group 2 (balls with biological motion), the forty videos of hand movements used in experimental group 1 were digitally manipulated so that the actor's hand was replaced with a ball, coloured using the same tones as the hand. The ball was the same size as the tip of the index finger that participants had to touch in experimental group 1 (30px. diameter) and was positioned at the same coordinates in each frame. An additional frame was created by positioning the ball mid-air before the first frame (where the ball contacts the table) creating an illusory "bounce" motion, providing a realistic context for the ball movement in order to reduce impressions of self-propelled movement that could also cue the observer that the motion is intentional (Luo & Baillargeon, 2005).

For Experimental group 3 (balls with non-biological motion), the forty videos from group 2 were digitally manipulated so that the ball now appeared to move in a straight line and at a constant speed after the bounce frame, eliminating the biological motion profile. To ensure that comparability of disappearance points between experimental groups, the line of best fit was calculated through the last four frames of each sequence of experimental group 1 (i.e. all possible disappearance points). The constant speed of the ball was created by recalculating the Y coordinates at equal distances along this line, between the first and last frame.

5.2.1.4 Procedure

An example trial sequence can be seen in Figure 5-1C. Participants completed four blocks of 48 trials in which each condition was presented an equal amount of times

(Straight/Efficient, Straight/Inefficient, Arched/Efficient, Arched/Inefficient). At the start of each trial, participants saw an instruction to “Hold the spacebar”, to which they pressed the spacebar with their right hand and kept it depressed. This ensured that they did not track the observed motion with their finger and could only initiate their response once the action sequence had disappeared. Participants then saw the first frame of the action sequence as a static image (the hand at rest in experimental group 1 and the “bouncing ball” frame in experimental groups 2 and 3) and were required to say “yes” into the microphone if there was an obstructing object present, and “no” if there was not.

The action sequence began 1000ms after a verbal response had been detected. Every third frame of the action sequence was presented for 80ms each, with a randomly selected sequence length of 5, 6, 7 or 8 frames (e.g. trials with a length of 8 frames showed frames 1-4-7-10-13-16-19-22). The final frame was then immediately replaced with the response frame, which showed the same scene without the moving object, creating the impression that the hand/ball had simply disappeared. Participants released the spacebar and, with their right hand, touched the screen where they thought the final position of the tip of the observed index finger was in group 1, or the final ball position in group 2 and 3. As soon as a response was registered, the next trial began.

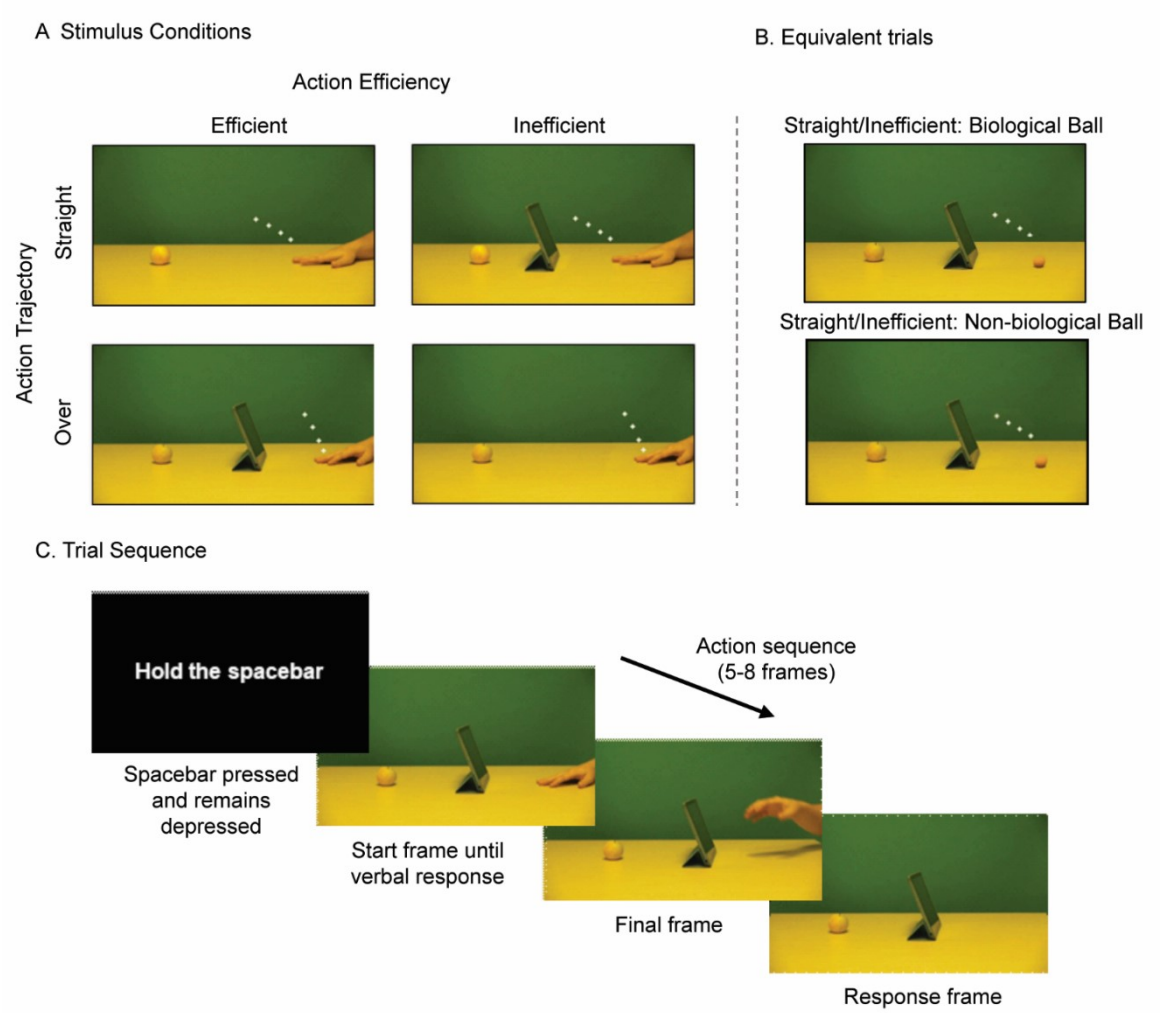


Figure 5-1. Experiments 3a-c. Stimulus conditions and trial sequence.

The stimulus conditions used in all three experimental groups are depicted in Panel A. The Action Trajectory was either straight or arched over. These Action Trajectories were either efficient or inefficient, governed by the presence or absence of an obstructing object. Panel B depicts an example of a Straight/Inefficient trial in the Biological Ball group (top) and the Non-Biological Ball group (bottom). Examples in Panel A depict the start frame and examples in Panel B depict the equivalent frame. The white markers depict the disappearance point of the index finger tip/ball in each of the four final frames. Panel C shows an example of a trial sequence in the Arched/Efficient condition of group 1. This trial sequence is equivalent across all experimental groups.

Note that the presentation of every third frame of the videos resulted in illusory "apparent" motion between the steps in the trajectory (Wertheimer, 1912). Such non-smooth motion retains the relevant characteristics of intentional biological motion

(e.g. parabolic path, bell-shaped velocity profile) and provides ideal conditions to measure predictive influences in motion perception, which are larger with apparent motion than smooth motion (Kerzel, 2003). This is in line with the notion that top-down influences that govern everyday perception become apparent the more the bottom-up sensory input becomes ambiguous or uncertain (e.g. through bi-stable images, Hohwy, Roepstorff, & Friston, 2008; visual noise, Gordon, Koenig-Robert, Tsuchiya, van Boxtel, & Hohwy, 2017). For motion, non-smooth step-wise presentation is assumed to disrupt low-level motion detectors, prompting a stronger weighting of top-down influences that compensate and "fill in" the intervening steps in the trajectory (Kok, Brouwer, van Gerven, & de Lange, 2013; Muckli et al., 2005; Yantis & Nakama, 1998).

5.2.2 Results

Data filtering was identical to our original experiment (Hudson et al., 2018). In all three experimental groups, trials were excluded if the correct response procedure was not followed (e.g. lifting the spacebar too early; 3.5%), or if response initiation or execution times were shorter than 200ms or more than 3SDs above the sample mean (2.2%, Initiation: mean =393.7ms, *SD*=173.3; Execution: mean =571.9ms, *SD*=203.3). Participants were excluded if too few trials remained after trial exclusions (< 50% valid trials, 3 participants), if the distance between the real and selected positions exceeded 3SD of the sample mean (mean =39.9 pixels, *SD*=18.9, 2 participants excluded), or if the correlation between the real and selected positions was more than 3SD below the median r value (X axis: median *r* =.940, *SD* = .041; Y axis: median *r* =.908, *SD* = .063, 4 participants excluded).

Analysis was conducted on the predictive perceptual bias by subtracting the real final coordinates of the tip of the index finger/ball from the participant's selected coordinates on each trial. This resulted in separate "difference" scores along the X and Y axis where positive X and Y scores represented a rightward and upward displacement respectively, and negative X and Y scores represented a leftward and downward displacement respectively. A score of 0 on both axes indicated that the participant selected the real final position exactly. These difference scores were entered into a 2x2x3 ANOVA for the X and Y axis separately, with Trajectory (straight vs arched) and Efficiency (efficient vs inefficient) as repeated-measures factors, and experimental group as a between-subjects factor.

The data from the original experiments, as well as further pilot studies in our lab, have shown that expectations of efficient action primarily induce biases on the Y-axis, but not the X-axis. This is consistent with the view that rather than viewing the current trajectory relative to the trajectory that was initially predicted (e.g. an arched trajectory when an obstacle was present), expectations of action efficiency reflect expectations about how the current trajectory will further develop. In other words, when seeing a straight reach towards an obstacle, one expects the hand to be merely lifted upwards to avoid the obstacle (rather than it being also displaced backwards to its corresponding location had it followed the arched trajectory from the outset). Similarly, when seeing an arched reach over empty space one expects the current reach would straighten downwards towards the goal object (rather than also being displaced forwards to where the hand would be had it followed the alternative straight trajectory). If the current results replicate this established pattern, displacement should therefore again primarily affect the Y axis (capturing this lifting

or lowering of the hand towards the target or away from the obstacle), but not the X-axis (indexing a displacement forwards/backwards to the alternative trajectory).

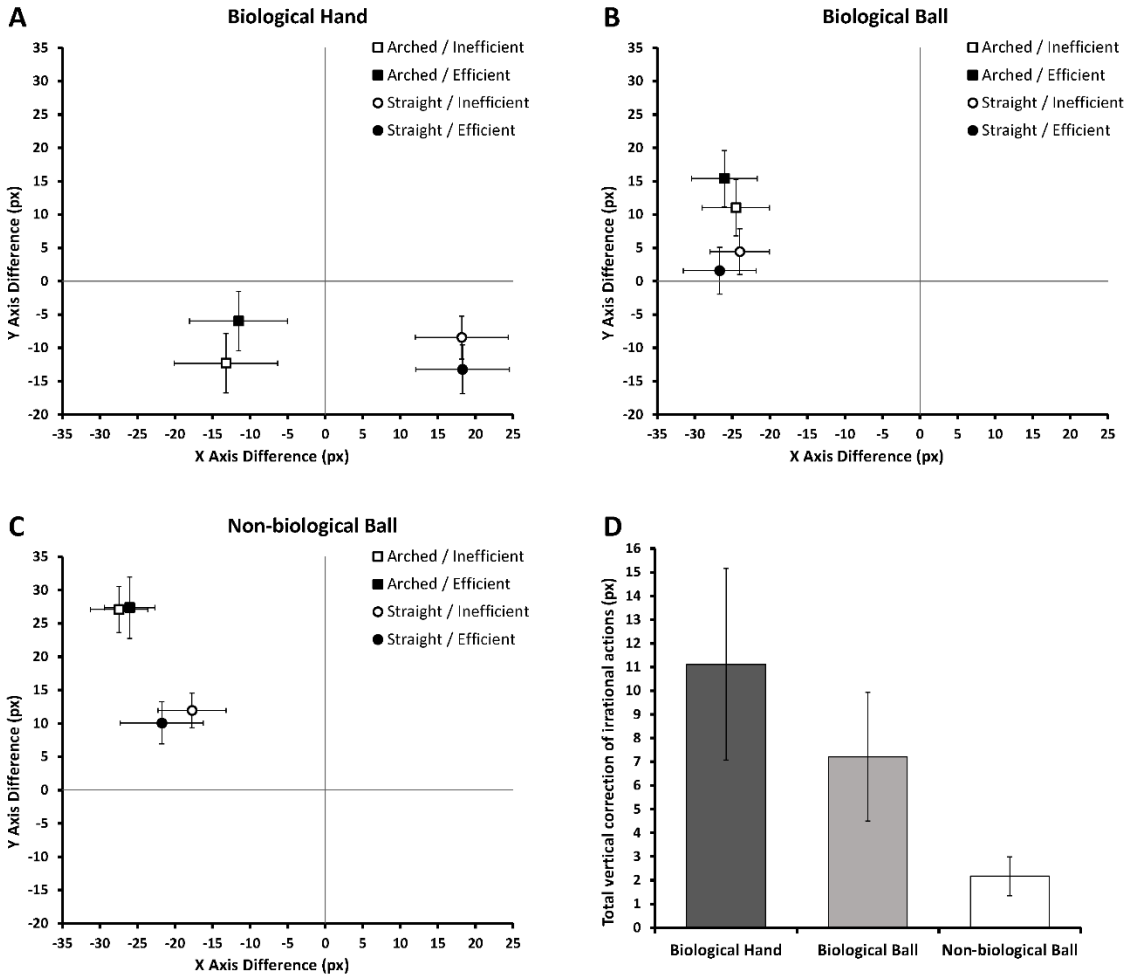


Figure 5-2. Experiments 3a-c. Results

The Trajectory X Efficiency interactions for the Biological Hand (A), Biological Ball (B), and Non-biological Ball (C) groups. The difference between the real final position and the selected final position is plotted for the X axis and Y axis. The real final position on any given trial is at point 0,0, as indicated on each plot. Panel D depicts a comparison of the size of the Y axis interaction in pixels, equivalent to the total amount by which inefficient actions were corrected towards a more efficient trajectory. Error bars depict 95% confidence intervals.

5.2.2.1 Y axis

If intentionality is perceptually instantiated, we predicted (1) that inefficient actions would be perceptually "corrected" towards the more efficient action alternative, and (2) that these biases should be strongest in experimental group 1 (hands with biological motion) but weaker when cues to intentionality are removed in groups 2 (balls with biological motion) and 3 (balls with non-biological motion). Indeed, the analysis revealed an interaction of Trajectory and Efficiency ($F(1,79) = 45.0$, $p < .001$, $\eta_p^2 = .363$), replicating our prior study (Hudson et al., 2018). Across groups, the disappearance points for straight trajectories were reported higher when the actions were inefficient (i.e. reaching towards an obstacle, 2.26px), than when the actions were efficient (no obstacle, -.967px; $t(81)=5.46$, $p<.001$, $d=.60$). Conversely, the perceived disappearance points for arched reaches were perceived to be lower for inefficient actions (7.87px) than for efficient actions (11.6px; $t(81)=4.81$, $p<.001$, $d=.53$).

Importantly, and as predicted, these biases differed between experimental groups, as indicated by an interaction of Trajectory, Efficiency and Experimental group ($F(1,79) = 6.47$, $p = .002$, $\eta_p^2 = .141$). Pairwise step-down comparisons showed that the interaction between Trajectory and Efficiency was smaller in the Non-biological Ball group (group 3) than in the Biological Hand group (group 1: $F(1,53)=11.7$, $p=.001$, $\eta_p^2=.181$), and the Biological Ball group ($F(1,51) =4.00$, $p=.051$, $\eta_p^2=.073$). No difference was found between Biological Hand group and the Biological Ball group (group 2: $F(1,54) =2.77$, $p=.102$, $\eta_p^2=.049$), although a Two One-Sided Tests (TOST) procedure (Lakens, 2017) indicated that the observed effect size ($d=.45$) was not significantly within the equivalence bounds of $\Delta L = –.53$ and $\Delta U = .53$,

$t(53.85) = -0.31$, $p = .38$ (equivalence bounds calculated as critical Cohen's d-values from our prior study investigating the same effect, Lakens, 2017). When experimental groups were analysed separately, the interaction between Trajectory and Efficiency was only present for the groups seeing Hands and Balls on biological motion trajectories (Biological Hand: $F(1,28) = 41.7$, $p < .001$, $\eta_p^2 = .598$; Biological Ball: $F(1,26) = 21.0$, $p < .001$, $\eta_p^2 = .447$), but not in the group viewing balls a non-biological motion trajectory (Non-biological Ball: $F(1,25) = 1.20$, $p = .284$, $\eta_p^2 = .046$). Indeed, the TOST procedure indicated that the observed effect size in the latter group (d=.21) was significantly within the equivalence bounds of ΔL = –.55 and ΔU =.55, $t(25) = -1.71$, $p = .05$.

As unpredicted effects are subject to alpha inflation in an ANOVA due to multiple testing (Cramer et al., 2015) all additional results in the analysis of Y-Axis and X-Axis should be interpreted with caution, and considered relative to a Bonferroni-adjusted alpha of .004. The analysis revealed an additional main effect of Trajectory that passed this threshold ($F(1,79) = 197.5$, $p < .001$, $\eta_p^2 = .714$), with perceived disappearance points of stimuli on arched trajectories being displaced further upwards (9.76px, $t(81)=9.72$, $p<.001$, d=1.1) than for straight trajectories (.67 px). This bias is consistent with the well-known predictive displacement in the direction of motion (e.g. further upwards for arched trajectories, but not for straight ones), known at the Representational Momentum effect (Hubbard, 2005). Interestingly, this forward displacement again differed between experimental groups, as indicated by an interaction of Trajectory and Experimental group that passed corrected thresholds ($F(1,79) = 40.4$, $p < .001$, $\eta_p^2 = .506$). Direct comparisons showed that the upwards displacements for arched trajectories were larger in Non-biological Ball group (group 3) than the Biological Ball group, (group 2: $F(1,51) = 12.9$, $p = .001$, $\eta_p^2 = .203$),

which in turn were larger than in the Biological Hand group, (group 1: $F(1,54) = 31.9$, $p < .001$, $\eta_p^2 = .371$). When analysing each experimental group separately, the upwards shift of straight trajectories was only present with ball stimuli, both when following biological, ($F(1,26) = 100.5$, $p < .001$, $\eta_p^2 = .794$), and non-biological trajectories, ($F(1,79) = 147.7$ $p < .001$, $\eta_p^2 = .855$), but not with moving hands, ($F(1,28) = 2.35$, $p = .136$, $\eta_p^2 = .077$). While not explicitly predicted, these displacements may reflect further changes to motion prediction depending on the presence of intentional cues. Balls, especially those that do not follow a biological motion trajectory, would be expected to continue on their upwards path, but hands would not when the goal of the reach is located towards the bottom, such as here. Nevertheless, due to the post-hoc nature of these findings, they should be treated with caution.

5.2.2.2 X axis

We did not have any prediction for the X axis. All effects are therefore subject to alpha inflation in an ANOVA and should be interpreted with caution, relative to a Bonferroni-adjusted alpha of .004 (Cramer et al., 2015). A main effect of Trajectory, ($F(1,79) = 199.0$, $p < .001$, $\eta_p^2 = .716$), passed this, which threshold was further qualified by an interaction of Trajectory and Experimental group ($F(1,79) = 112.9$, $p < .001$, $\eta_p^2 = .741$). As can be seen in Figure 5-2, perceptual judgments of hands – but not balls – on arched trajectories were generally biased leftwards and rightwards for hands on straight trajectories. This difference replicates previous results and simply reflects stimulus differences between the hand shapes of the naturally recorded reaches on straight and arched trajectories, specifically the further rightwards centre of gravity for hands on straight trajectories, which biases location judgments (Coren

& Hoenig, 1972). No other effects passed the Bonferroni-adjusted thresholds of .004. Specifically, there was no main effect of Efficiency ($F(1,79) = .4.66$, $p = .034$, $\eta_p^2 = .056$), no interaction between Efficiency and Experimental group ($F(1,79) = 5.42$, $p = .006$, $\eta_p^2 = .121$), no interaction between Trajectory and Efficiency ($F(1,79) = 5.15$, $p = .026$, $\eta_p^2 = .061$) and no three-way interaction between Trajectory, Efficiency and Experimental group ($F(1,79) = 1.25$, $p = .293$, $\eta_p^2 = .031$).

5.2.2.3 Testing for general differences in attention between groups

In an exploratory analysis, we tested whether the observed differences between groups can be explained by more general differences in attention towards the biological and non-biological stimuli. In particular, it is well-established that agentive stimuli with a biological motion profile attract attention (Guerrero & Calvillo, 2016; Lindemann, Nuku, Rueschemeyer, & Bekkering, 2011; Pratt, Radulescu, Guo, & Abrams, 2010). To ensure that our results cannot be explained simply by more attentive perception of the more biological stimuli, we used the across-trial correlations between actual disappearance points and participants' judgments that we used to identify participants that did not follow the task (i.e. if the reported x coordinates did not bear enough relationship to the actual coordinates). If participants attend more strongly to biological stimuli than to non-biological stimuli, one would expect their judgements to be more accurate and to more closely follow what was observed, resulting in smaller deviations for biological hand stimuli than for the other, less intentional stimulus types. We found no evidence for this prediction. While these correlations were generally high, they were, if anything, higher in the ball conditions in which participants' judgments are less affected by their expectations (Hands, mean x $r = .91$, mean y $r = .88$; biological ball, mean x $r = .95$, mean y $r =$

.93; non-biological ball, mean x r = .92, mean y r = .90). While this runs counter to the argument for decreased attention in the non-biological conditions, it is fully in line with our proposal of a stronger reliance on prior expectations as soon as intentions can be attributed to these stimuli. Indeed, as predicted from this hypothesis, participants' across-trial correlations between actual and selected coordinates correlated negatively with how much they are affected by their expectations (r=-.30, p=.006), even when gross-between groups differences are factored out via z standardization in each group (r=-.31, p=.005). Thus, across all participants in the three groups, differences in the ability to track the actual disappearance points show, if anything, better accuracy in the non-biological groups, which can be explained by an (over-) reliance on prior expectations for stimuli that provide intentional cues.

5.2.3 Discussion

Previous studies have shown that perceptual representations of observed actions are predictively biased towards the goals attributed to them (Hudson et al., 2017; Hudson et al., 2012; Hudson & Jellema, 2011; Hudson et al., 2009; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016) and that these predictions are informed by the assumption of efficient action, reflecting the specific trajectories that would allow an actor to efficiently reach the inferred goal (Hudson et al., 2018). To investigate if these expectations emerge from assumptions about action intentionality, we asked participants to watch moving stimuli and to accurately report the object's last seen position after it suddenly disappeared. We tested whether perceptual reports would again be predictively biased towards the expected trajectory (Hudson et al., 2017; Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016) but varied whether the

stimulus was a hand with biological motion kinematics (i.e. bell-shaped velocity profile of reaching, Beggs & Howarth, 1972), a non-agentive ball that travelled the same biological motion trajectory as the hand, or a ball travelling a non-biological trajectory.

Replicating our prior studies, perceptual reports of hand disappearance points were not veridical, but “corrected” towards the expected action kinematics of a rational, efficient actor. The perceived disappearance points of hands reaching straight towards an obstacle were reported higher than if the path was clear. Similarly, the perceived disappearance point of arched reaches was perceived lower if there was no obstacle to reach across, compared to when there was an obstacle. Importantly, our new data now show that these biases towards efficient action depend on cues to intentionality. The biases were numerically reduced when participants watched a non-intentional object – a ball – travel on the same biological motion trajectory, starting slowly and speeding up along, as if self-propelled. They were almost completely eliminated, when the same ball was now seen travelling with a non-biological trajectory that nevertheless traversed, on average, the same path of motion as the hands, but did not show the characteristic bell-shaped velocity profile of goal-directed reaches (Beggs & Howarth, 1972).

These results confirm first that, as in our prior studies, observers predict the ideal action trajectory a rational actor would take that is fully aware of all relevant environmental constraints. Second, they show that these predictions influenced the perceptual judgments of observed actions, subtly biasing them towards the most efficient trajectory. These findings are therefore in line with predictive processing models of social perception (Bach et al., 2014; Bach & Schenke, 2017; Hudson,

Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016; Kilner et al., 2007a, 2007b; Zaki, 2013), which assume that the perceptual experience of others' actions emerges from an integration of bottom-up sensory information and prior assumptions about others' goals and how they would (best) realise them. Our data now show, third, that when observing the behaviour of others these predictions of efficient action depend on bottom-up cues to intentionality derived from the objects' semantics and its trajectory and motion profile. Both types of cues have been previously identified as a basis for attributing intentionality to observed agents in children (Baron-Cohen, 1997; Leslie, 1994; Morewedge et al., 2007; Rakison & Poulin-Dubois, 2001). The finding that these cues also modulate predictive biases towards efficient action in adult action observation directly supports the proposal that these predictions emerge from the attribution of intention to the observed actions (Hudson et al., 2017; Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016), which then inform their perceptual representation.

During everyday action observation these top-down influences can fulfil several important functions. First, they can disambiguate perception by compensating for the perceptual "blurring" during motion perception (i.e. motion sharpening, Bex et al., 1995; Hammett, 1997), or filling in missing steps of the input (Muckli et al., 2005). Second they can support planning of one's own actions, allowing them to be coordinated with the others' future behaviour or the end-state of their actions (Sebanz, Bekkering, & Knoblich, 2006). Finally, they can be compared to actual behaviour, triggering revisions of prior assumptions if prediction errors become too large (Clark, 2013; Hohwy, 2013), signalling, for example, that a behaviour may not be intentional after all, or that the actor is not aware of all relevant environmental

constraints (e.g., they may not have seen the obstacle). As such, they may underlie the proposed link between teleological perception of others' behaviour and more sophisticated theory of mind and mentalizing processes (Csibra & Gergely, 2007).

Further work now needs to resolve via which mechanisms cues of intentionality induce the predictive biases towards efficient action. One possibility is that the biases emerge via predictive mechanisms in one's own motor system (Ansuini, Cavallo, Bertone, & Becchio, 2015; Csibra, 2008; Kilner et al., 2007a, 2007b; Otten, Seth, & Pinto, 2017). On such views, people make higher-level "cognitive" attributions of intentions of others and then feed these goals into their own motor system to predict the kinematics they would need to achieve if they were in the actor's place. Indeed, the perceptual effects observed here bear a striking similarity to similar motoric effects that can be measured when people watch others' behaviour. Both behavioural and neuroimaging studies suggest that, during action observation, one's own motor system does not only mirror the actually seen behaviour (e.g. a finger being depressed) but also the behaviour that is only *predicted* from the goals attributed to the actor, even if it is not actually observed (Bach et al., 2011; Bach et al., 2014, e.g., finger held up by a clamp, Liepelt, Von Cramon, & Brass, 2008). Even if one watches an inanimate ball that one has experience of controlling oneself, one's motor behaviour subtly captures both the ball's actual trajectory and the trajectory one intended for it to travel on (De Maeght & Prinz, 2004; Knuf, Aschersleben, & Prinz, 2001).These motoric changes might therefore index the recruitment of such predictive (forward modelling) mechanisms that have evolved for the control of one's own actions but are applied to the actions of others.

An alternative possibility is that attributions of intentionality are made within the (higher-level) perceptual system. It is well-known that the perceptual system itself can make sophisticated “unconscious inferences” about objects, extracting, for example, the real colour of a stimulus by subtracting out cues to shading and illumination (Bloj, Kersten, & Hurlbert, 1999). In the same way, the perceptual system could use object and motion information (e.g., balls vs. hands; biological vs. non-biological motion profiles) to make inferences about the intentionality of a moving object (Scholl & Gao, 2013). Indeed, several imaging studies suggest that such cues to intentionality act on lower-level regions within higher-level visual cortex, such as the superior temporal sulcus (Grossman et al., 2000; Saygin, 2007). Moreover, it well known that children can attribute intentionality to stimuli which are unlikely to engender motor activation, such as abstract geometric shapes or biomechanically impossible actions (Heider & Simmel, 1944; Southgate, Johnson, & Csibra, 2008) or that they process action efficiency before they have competence in the observed action (Gredebäck & Melinder, 2010; Sodian, Schoeppner, & Metz, 2004). Local interactions within the perceptual system could explain such observation. In such views, the motoric activation measured during action observation described above therefore does not reflect the origin of the perceptual effects, but a mere passive “motor resonance” that captures instead the changes to the action’s perceptual representation that has already occurred.

While we are sympathetic to both explanations (Bach & Schenke, 2017), and we do not deem them as mutually exclusive, our prior data seems to be more consistent with the latter, perceptual locus of effects. In our original study (Hudson, Nicholson, Simpson, et al., 2016), we observed that while attributing goals (e.g. to reach or withdraw) to others reliably biased perceptual measures towards these goals, the

same was not true for when these action possibilities were motorically activated (i.e. by asking participants to make a forward or backwards movement with their own hand). While this conclusion is certainly preliminary, and needs to be supported by further studies, it makes a strong causal role of motoric processes unlikely.

Another question is how the present effects on perceptual judgments emerge. Several studies, both psychophysical and based on neuroimaging, have shown that predictions can exert downstream effects on early perceptual processes, across different modalities (e.g., vision, Ekman et al., 2017; Muckli et al., 2005, audition, Kondo et al., 2017), providing sensory “templates” of expected stimulation (Ekman et al., 2017), or filling in missing information during apparent motion (Avenanti et al., 2012; Muckli et al., 2005). Others, however, argue that expectations influence primarily decision-related processes that integrate bottom-up with top-down information on all levels of the hierarchy (Bang & Rahnev, 2017; Rungratsameetaweemana, Itthipuripat, Salazar, & Serences, 2018), or that they reflect attentional modulations of the response properties of neurons in early sensory areas (Desimone & Duncan, 1995; Serences & Kastner, 2014). Others argue that many of the psychophysical effects of expectation may in fact reflect testing artefacts or demand effects, when participants realise what is being tested (Durgin et al., 2009; Firestone & Scholl, 2016).

While the precise mechanism has to be confirmed, several aspects of prior studies (Hudson et al., 2018) imply a role in the action’s perceptual representation. First, when asked during piloting of the original set of studies, participants were unaware of the experimental hypotheses, arguing against demand effects. Second, the effects were present already very briefly (250 ms) after action offset, in psychophysical

probe judgment tasks (Hudson et al., 2017; Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016, for a review of similar findings in non-biological motion perception, see Hubbard, 2015) that has been shown to be relatively robust against cognitive control processes (Courtney & Hubbard, 2008; Ruppel, Fleming, & Hubbard, 2009). Third, and most importantly, the biases towards efficient action were disrupted by brief (560 ms.) dynamic visual noise masks that interfere with the re-entrant feedback from higher cortical areas with visual cortex that is required for the stabilisation of percepts for conscious access, during both perception (Breitmeyer & Öğmen, 2006; Fahrenfort et al., 2007; Kinsbourne & Warrington, 1962; Lamme et al., 2002) and imagery (Dijkstra, Mostert, de Lange, Bosch, & van Gerven, 2018). The observed biases in perceptual judgments are therefore unlikely to stem from unspecific perceptual changes in memory or motor control (Firestone & Scholl, 2016; for an example for perceptual changes in action memory, see, Ianì, Mazzoni, & Bucciarelli, 2018). Instead, we propose that they either play a role in ongoing motion perception emerging from the re-current interactions between lower and higher visual regions involved in stabilising percepts and compensating for the substantial blurring during motion perception.

5.2.4 Conclusions

The principle of efficient action allows observers to predict ideal reference trajectories that intentional actions will follow, given that the agent is fully aware of all relevant environmental constraints. The data presented here confirm that these predictions are at least partially perceptually represented and influence perceptual judgments of others actions, biasing them towards these expectations. They show that these predictions emerge from attributions of intentionality to the observed actor,

triggered by the perception of biological “agentive” objects and kinematics that follow biological motion profiles.

5.2.5 References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7*(10), 1057-1058. doi:10.1038/nn1312

Ansuini, C., Cavallo, A., Bertone, C., & Becchio, C. (2015). Intentions in the Brain. *The Neuroscientist, 21*(2), 126-135. doi:10.1177/1073858414533827

Avenanti, A., Annella, L., Candidi, M., Urgesi, C., & Aglioti, S. M. (2012). Compensatory plasticity in the action observation network: virtual lesions of STS enhance anticipatory simulation of seen actions. *Cerebral cortex, 23*(3), 570-580.

Bach, P., Bayliss, A. P., & Tipper, S. P. (2011). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychonomic Bulletin & Review, 18*(1), 171-176. doi:10.3758/s13423-010-0029-x

Bach, P., Knoblich, G., Gunter, T. C., Friederici, A. D., & Prinz, W. (2005). Action Comprehension: Deriving Spatial and Functional Relations. *Journal of Experimental Psychology: Human Perception and Performance, 31*(3), 465-479. doi:10.1037/0096-1523.31.3.465

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Frontiers in Human Neuroscience, 8*. doi:10.3389/fnhum.2014.00254

Bach, P., Nicholson, T., & Hudson, M. (2015). Pattern completion does not negate matching: a response to Uithol and Maranesi. . *Frontiers in Human Neuroscience, 9*. doi:10.3389/fnhum.2015.00685

Bach, P., & Schenke, K. C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass, 11*(7), e12312. doi:10.1111/spc3.12312

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological Reasoning in Infancy. *Annual Review of Psychology, 67*(1), 159-186. doi:10.1146/annurev-psych-010213-115033

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*(3), 329-349. doi:10.1016/j.cognition.2009.07.005

Bang, J. W., & Rahnev, D. (2017). Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Scientific Reports, 7*(1). doi:10.1038/s41598-017-16885-2

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. Cambridge, Massachusetts: MIT press.

Beggs, W. D. A., & Howarth, C. I. (1972). The Movement of the Hand towards a Target. *Quarterly Journal of Experimental Psychology, 24*(4), 448-453. doi:10.1080/14640747208400304

Bex, P. J., Edgar, G. K., & Smith, A. T. (1995). Sharpening of drifting, blurred images. *Vision Research, 35*(18), 2539-2546. doi:10.1016/0042-6989(95)00060-d

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature, 402*(6764), 877-879. doi:10.1038/47245

Breitmeyer, B., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision*: Oxford University Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(03), 181-204. doi:10.1017/s0140525x12000477

Coren, S., & Hoenig, P. (1972). Effect of Non-Target Stimuli upon Length of Voluntary Saccades. *Perceptual and Motor Skills, 34*(2), 499-508. doi:10.2466/pms.1972.34.2.499

Courtney, J. R., & Hubbard, T. L. (2008). Spatial Memory and Explicit Knowledge: An Effect of Instruction on Representational Momentum. *Quarterly Journal of*

Experimental Psychology, 61(12), 1778-1784. doi:10.1080/17470210802194217

Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., . . . Wagenmakers, E.-J. (2015). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review, 23*(2), 640-647. doi:10.3758/s13423-015-0913-5

Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorymotor foundations of higher cognition. Attention and performance XXII* (pp. 435-459). Oxford: Oxford University Press.

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica, 124*(1), 60-78. doi:10.1016/j.actpsy.2006.09.007

De Maeght, S., & Prinz, W. (2004). Action induction through action observation. *Psychological Research, 68*(2-3), 97-114.

Dennett, D. C. (1987). *The intentional stance.* Cambridge, MA: MIT Press.

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience, 18*(1), 193-222. doi:10.1146/annurev.neuro.18.1.193

Dijkstra, N., Mostert, P., de Lange, F., Bosch, S. E., & van Gerven, M. (2018). Differential temporal dynamics during visual imagery and perception. In: eLife.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review, 16*(5), 964-969. doi:10.3758/pbr.16.5.964

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications, 8*, 15276. doi:10.1038/ncomms15276

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2007). Masking Disrupts Reentrant Processing in Human Visual Cortex. *Journal of Cognitive Neuroscience, 19*(9), 1488-1497. doi:10.1162/jocn.2007.19.9.1488

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*(7), 878-879. doi:10.1038/nn1729

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences, 39*. doi:10.1017/s0140525x15000965

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(1), 126-132. doi:10.1037//0278-7393.10.1.126

Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks, 22*(8), 1093-1104. doi:10.1016/j.neunet.2009.07.023

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature, 415*(6873), 755-755. doi:10.1038/415755a

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends in Cognitive Sciences, 7*(7), 287-292. doi:10.1016/s1364-6613(03)00128-1

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165-193. doi:10.1016/0010-0277(95)00661-h

Gordon, N., Koenig-Robert, R., Tsuchiya, N., van Boxtel, J. J., & Hohwy, J. (2017). Neural markers of predictive coding under perceptual uncertainty revealed with Hierarchical Frequency Tagging. *Elife, 6*, e22749.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114*(2), 197-206.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain Areas Involved in Perception of Biological Motion. *Journal of Cognitive Neuroscience, 12*(5), 711-720. doi:10.1162/089892900562417

Guerrero, G., & Calvillo, D. P. (2016). Animacy increases second target reporting in a rapid serial visual presentation task. *Psychonomic Bulletin & Review, 23*(6), 1832-1838.

Hammett, S. T. (1997). Motion blur and motion sharpening in the human visual system. *Vision Research, 37*(18), 2505-2510. doi:10.1016/s0042-6989(97)00059-x

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American journal of psychology, 57*(2), 243-259.

Hohwy, J. (2013). *The Predictive Mind*: Oxford University Press.

Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition, 108*(3), 687-701.

Hubbard, T. L. (1995). Cognitive representation of motion: Evidence for friction and gravity analogues. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 241-254. doi:10.1037//0278-7393.21.1.241

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12*(5), 822-851. doi:10.3758/bf03196775

Hubbard, T. L. (2015). The varieties of momentum-like experience. *Psychological Bulletin, 141*(6), 1081-1119. doi:10.1037/bul0000016

Hudson, M., Bach, P., & Nicholson, T. (2017). You said you would! The predictability of other's behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance, 44*(2), 320-335. doi:10.1037/xhp0000451

Hudson, M., Burnett, H. G., & Jellema, T. (2012). Anticipation of Action Intentions in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 42*(8), 1684-1693. doi:10.1007/s10803-011-1410-y

Hudson, M., & Jellema, T. (2011). Resolving ambiguous behavioral intentions by means of involuntary prioritization of gaze processing. *Emotion, 11*(3), 681-686. doi:10.1037/a0023264

Hudson, M., Liu, C. H., & Jellema, T. (2009). Anticipating intentional actions: The effect of eye gaze direction on the judgment of head rotation. *Cognition, 112*(3), 423-434. doi:10.1016/j.cognition.2009.06.011

Hudson, M., McDonough, K. L., Edwards, R., & Bach, P. (2018). Perceptual Teleology: Expectations of Action Efficiency Bias Social Perception.

Proceedings of the Royal Society: B.(285(1884)), 20180638. doi:10.1098/rspb.2018.0638

Hudson, M., Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition, 146*, 245-250. doi:10.1016/j.cognition.2015.09.021

Hudson, M., Nicholson, T., Simpson, W. A., Ellis, R., & Bach, P. (2016). One step ahead: The perceived kinematics of others' actions are biased toward expected goals. *Journal of Experimental Psychology: General, 145*(1), 1-7. doi:10.1037/xge0000126

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1644), 20130490-20130490. doi:10.1098/rstb.2013.0490

Ianì, F., Mazzoni, G., & Bucciarelli, M. (2018). The role of kinematic mental simulation in creating false memories. *Journal of Cognitive Psychology, 30*(3), 292-306.

Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences, 4*(1), 22-28. doi:10.1016/s1364-6613(99)01414-x

Johnson, S. C. (2003). Detecting agents. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1431), 549-559. doi:10.1098/rstb.2002.1237

Jordan, J. S., & Hunsinger, M. (2008). Learned patterns of action-effect anticipation contribute to the spatial displacement of continuously moving stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 34*(1), 113-124. doi:10.1037/0096-1523.34.1.113

Kerzel, D. (2003). Mental extrapolation of target position is strongest with weak motion signals and motor responses. *Vision Research, 43*(25), 2623-2635.

Kessler, K., Gordon, L., Cessford, K., & Lages, M. (2010). Characteristics of motor resonance predict the pattern of flash-lag effects for biological motion. *PLOS ONE, 5*(1), e8258.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007a). The mirror-neuron system: a Bayesian perspective. *NeuroReport, 18*(6), 619-623. doi:10.1097/wnr.0b013e3281139ed0

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007b). Predictive coding: an account of the mirror neuron system. *Cognitive Processing, 8*(3), 159-166. doi:10.1007/s10339-007-0170-2

Kinsbourne, M., & Warrington, E. K. (1962). The Effect of an After-coming Random Pattern on the Perception of Brief Visual Stimuli. *Quarterly Journal of Experimental Psychology, 14*(4), 223-234. doi:10.1080/17470216208416540

Knuf, L., Aschersleben, G., & Prinz, W. (2001). An analysis of ideomotor action. *Journal of Experimental Psychology: General, 130*(4), 779.

Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience, 33*(41), 16275-16284.

Kondo, H. M., Farkas, D., Denham, S. L., Asai, T., & Winkler, I. (2017). Auditory multistability and neurotransmitter concentrations in the human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1714), 20160110. doi:10.1098/rstb.2016.0110

Lakens, D. (2017). Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*(4), 355-362.

Lamme, V. A. F., Zipser, K., & Spekreijse, H. (2002). Masking Interrupts Figure-Ground Signals in V1. *Journal of Cognitive Neuroscience, 14*(7), 1044-1053. doi:10.1162/089892902320474490

Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain*

specificity in cognition and culture (pp. 119-148). New York: Cambridge University Press.

Liepelt, R., Von Cramon, D. Y., & Brass, M. (2008). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *Neuroimage, 43*(4), 784-792.

Lindemann, O., Nuku, P., Rueschemeyer, S.-A., & Bekkering, H. (2011). Grasping the other's attention: The role of animacy in action cueing of joint attention. *Vision Research, 51*(8), 940-944.

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition, 160*, 35-42. doi:10.1016/j.cognition.2016.12.007

Luo, Y., & Baillargeon, R. (2005). Can a Self-Propelled Box Have a Goal?: Psychological Reasoning in 5-Month-Old Infants. *Psychological Science, 16*(8), 601-608. doi:10.1111/j.1467-9280.2005.01582.x

Morewedge, C. K., Preston, J., & Wegner, D. M. (2007). Timescale bias in the attribution of mind. *Journal of Personality and Social Psychology, 93*(1), 1-11. doi:10.1037/0022-3514.93.1.1

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary Visual Cortex Activity along the Apparent-Motion Trace Reflects Illusory Perception. *PLoS Biology, 3*(8), e265. doi:10.1371/journal.pbio.0030265

Otten, M., Seth, A. K., & Pinto, Y. (2017). A social Bayesian brain: How social knowledge can shape visual perception. *Brain and Cognition, 112*, 69-77. doi:10.1016/j.bandc.2016.05.002

Pozzo, T., Papaxanthis, C., Petit, J. L., Schweighofer, N., & Stucchi, N. (2006). Kinematic features of movement tunes perception and action coupling. *Behavioural Brain Research, 169*(1), 75-82. doi:10.1016/j.bbr.2005.12.005

Pratt, J., Radulescu, P. V., Guo, R. M., & Abrams, R. A. (2010). It's alive! Animate motion captures visual attention. *Psychological Science, 21*(11), 1724-1730.

Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate-inanimate distinction. *Psychological Bulletin, 127*(2), 209-228. doi:10.1037//0033-2909.127.2.209

Rochat, M. J., Serra, E., Fadiga, L., & Gallese, V. (2008). The Evolution of Social Cognition: Goal Familiarity Shapes Monkeys' Action Understanding. *Current Biology, 18*(3), 227-232. doi:10.1016/j.cub.2007.12.021

Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A., & Serences, J. T. (2018). Expectations do not alter early sensory processing during perceptual decision

making. *The Journal of Neuroscience*, 3638-3617. doi:10.1523/jneurosci.3638-17.2018

Ruppel, S. E., Fleming, C. N., & Hubbard, T. L. (2009). Representational momentum is not (totally) impervious to error feedback. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 63*(1), 49-58. doi:10.1037/a0013980

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition, 119*(2), 242-252. doi:10.1016/j.cognition.2011.01.014

Saunier, G., Papaxanthis, C., Vargas, C. D., & Pozzo, T. (2008). Inference of complex human motion requires internal models of action: behavioral evidence. *Experimental Brain Research, 185*(3), 399-409. doi:10.1007/s00221-007-1162-2

Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain, 130*(9), 2452-2461. doi:10.1093/brain/awm162

Schlaffke, L., Golisch, A., Haag, L. M., Lenz, M., Heba, S., Lissek, S., . . . Tegenthoff, M. (2015). The brain's dress code: How The Dress allows to decode the neuronal pathway of an optical illusion. *Cortex, 73*, 271-275. doi:10.1016/j.cortex.2015.08.017

Scholl, B. J., & Gao, T. (2013). Perceiving Animacy and Intentionality: Visual processing or higher-level judgment. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social Perception: Detection and interpretation of animacy, agency, and intention,* (pp. 197-230): The MIT Press.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences, 10*(2), 70-76. doi:10.1016/j.tics.2005.12.009

Serences, J. T., & Kastner, S. (2014). A Multi-level Account of Selective Attention. In *Oxford Handbooks Online*: Oxford University Press.

Shiffrar, M., & Freyd, J. J. (1993). Timing and Apparent Motion Path Choice With Human Body Photographs. *Psychological Science, 4*(6), 379-384. doi:10.1111/j.1467-9280.1993.tb00585.x

Sodian, B., Schoeppner, B., & Metz, U. (2004). Do infants apply the principle of rational action to human agents? *Infant Behavior and Development, 27*(1), 31-41.

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition, 107*(3), 1059-1069.

Wellman, H. M., & Brandone, A. C. (2009). Early intention understandings that are common to primates predict children's later theory of mind. *Current Opinion in Neurobiology, 19*(1), 57-62. doi:10.1016/j.conb.2009.02.004

Wertheimer, M. (1912). Experimentelle studien uber das sehen von bewegung. *Zeitschrift fur Psychologie, 61*.

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience, 1*(6), 508-512. doi:10.1038/2226

Zaki, J. (2013). Cue integration: A common framework for social cognition and physical perception. *Perspectives on Psychological Science, 8*(3), 296-312. doi:10.1177/1745691613475454

6 Chapter Six - Online action prediction

All experiments so far in thesis have revealed a reliance upon expectations of action efficiency to guide the perception of others actions. These expectations rest upon attributions of intentionality, triggered by semantic and motion cues from the observed action. They provide a perceptual reference signal for the specific trajectory a rational and intentional actor would make through a given scene to achieve their goal, which is constantly tested against the perceptual input, inducing a subtle perceptual confirmation bias that distorts observed action kinematics towards expectations. In all these experiments, the goals and intentions of the actor (to reach the target object) and the environmental constraints (the presence or absence of an obstacle) were known prior to action onset. In the real world, the mental states of others (i.e. their goals, intentions, knowledge of the world) are not explicitly available beforehand, and instead have be derived dynamically from their ongoing behaviour from cues embedded in their action kinematics.

The experiments in Chapter Six investigated whether predictions of others upcoming actions are not only derived before action onset, but are dynamically adjusted “on-line” as more goal and intention information becomes available from the unfolding kinematics and then integrated with the action’s perceptual representation. Such findings would provide applications of the current findings to real world social interactions, revealing how humans so effortlessly attribute meaning to others behaviours in such a complex and dynamic social world.

The experiments in this chapter are currently under review for publication and have been uploaded to the preprint server, PsyArX. They are presented in their submitted format (green copy).

6.1 Affordance matching predictively shapes the perceptual representation of others' ongoing actions.

Katrina L. McDonough, Marcello Costantini, Matthew Hudson, Patric Bach.

Abstract

Predictive processing accounts of social perception argue that action observation is a predictive process, in which inferences about others goals are constantly tested against the perceptual input, inducing a subtle perceptual confirmation bias that distorts observed action kinematics towards the inferred goals. Here we test whether such biases are induced even when goals are not explicitly given but have to be derived from the unfolding action kinematics. In two experiments, participants briefly saw an actor reach ambiguously towards a large object and a small object, with either a whole-hand power grip or a precision grip. During its course, the hand suddenly disappeared, and participants reported its last seen position on a touch-screen. As predicted, judgments were consistently biased towards apparent action targets, such that power grips were perceived closer to large objects and precision grips closer to small objects, even if the reach kinematics were identical. Strikingly, these biases were independent of participants' explicit goal judgments. They were of equal size when action goals had to be explicitly derived in each trial (Experiment 4a) or not (Experiment 4b) and, across trials and across participants, explicit judgments and perceptual biases were uncorrelated. This provides evidence, for the first time, that people make on-line adjustments of observed actions based on the match between hand grip and object goals, distorting their perceptual representation towards implied goals. These distortions may not reflect high-level goal assumptions, but emerge from relatively low-level processing of kinematic features within the perceptual system.

6.2 Experiments 4a and 4b

The ability to understand and predict other people's behaviour is a cornerstone of human social cognition and makes people's sophisticated interactions with others possible. A parent constantly monitors their child's goals, and intervenes when it reaches for the hot cup of coffee instead of the toy right next to it. In sports, players foresee each other's behaviour, fluently passing a ball to a team mate's future position. In contrast, deficits in the ability to understand others' behaviour are a hallmark of several conditions that bring with them marked impairments in social interactions, such as autism (Koster-Hale & Saxe, 2013; Pellicano & Burr, 2012; von der Lühe et al., 2016).

These abilities for social perception are conventionally conceptualized as a simple bottom-up process, in which incoming visual information about others' behaviour is matched to one's higher-level motor – or conceptual – knowledge about it, so that the action's meaning and associated mental states can be derived (Gallese & Goldman, 1998; Rizzolatti & Sinigaglia, 2010; Rizzolatti, Cattaneo, Fabbri-Destro & Rozzi, 2014). However, there is no one-to-one mapping between stimuli and meaning that such a mechanism could rely on (Ansuini, Cavallo, Bertone, & Becchio, 2015; Jacob & Jeannerod, 2005; Uithol, van Rooij, Bekkering, & Haselager, 2011). The same behaviour can mean multiple things in different contexts (e.g. a smile), and the same goals can be accomplished by multiple behaviours (e.g., closing a drawer with one's hand vs hip). It has therefore been argued that social perception is better understood as a *predictive* process in which the brain constantly tests hypotheses about the observed action against the perceptual input (Bach, Nicholson, & Hudson, 2014; Bach & Schenke, 2017; Csibra, 2008; Donnarumma, Costantini,

Ambrosini, Friston, & Pezzulo, 2017; Kilner, Friston, & Frith, 2007[ab]). In such accounts, any assumption about others' goals and beliefs – derived perhaps from prior knowledge about the individual (Joyce, Schenke, Bayliss, & Bach, 2015; Schenke, Wyer, & Bach, 2016) or from contextual information (e.g., objects, Bach, Knoblich, Gunter, Friederici, & Prinz, 2005; Jacquet et al, 2012; Kalénine et al., 2013, Nicholson, Roser, & Bach, 2017; gaze and emotional expressions, Adams, Ambady, Macrae, & Kleck, 2006; Frischen & Tipper, 2006) – is translated into predictions about which behaviour should be observed if these assumptions were correct and is superimposed over the perceptual input. Such an integration would not only help to stabilize perception, filling in gaps in the input (e.g. in the case of occlusion, Prinz & Rapinett, 2008) or compensate for the considerable noise during motion perception (Hammett, 1997), but would also let mismatching behaviour of others' stand out, so that our assumptions about them can be revised until they better explain their behaviour.

We recently developed an experimental paradigm that can make these predictions visible (Hudson, Bach, & Nicholson, 2017; Hudson, McDonough, Edwards, & Bach, 2018; Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016; McDonough, Hudson, & Bach, 2019). This paradigm rests on the assumption that, if predictions indeed act on perceptual representations (Bar, 2004; de Lange, Heilbron, & Kok, 2018; Ekman, Kok, & de Lange, 2017) then every prediction one makes about another person may subtly bias the perception of their forthcoming actions, especially in case of uncertainty, such as the visual blurring during motion perception. Thus, in the same way as prior expectations in the non-social world cause us to see a colour differently (Bloj, Kersten, & Hurlbert, 1999; see for an application to the blue/gold dress illusions, Chetverikov & Ivanchei, 2016) or

shapes as either convex or concave based on the surrounding illumination (Adams, Graf, & Ernst, 2004), our prior knowledge of other people – their goals and intentions – may subtly shape the perceptual experience of their actions.

This is indeed what we observed. Participants heard an actor make a statement about their goal – “I’ll take it” or “I’ll leave it” – before they briefly saw him start to reach for an object or withdraw from it. The action disappeared mid-motion and participants indicated the perceived vanishing point, either by comparing it to a probe stimulus shortly after stimulus offset (Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016) or by indicating its disappearance point on a touch screen (Hudson et al., 2017; Hudson et al., 2018). The results consistently revealed predictive biases on these perceptual judgments. First, hands were generally reported to have disappeared further along the trajectory than what was actually seen, capturing lower-level predictions based on the action’s prior course (e.g. Representational Momentum, Freyd & Finke, 1984; Hubbard, 2005). Second, and more importantly, they revealed an influence of goals attributed to the actor: hands were reported to have disappeared further towards the object when the actor said they would take it and further away from the object when the actor said they wanted to leave it. Other studies extended these findings, showing that similar distortions can be induced when the participant instructed the observed actor (Hudson, Nicholson, Ellis, et al., 2016) and that the observed actor’s long-term reliability to do as they said modulates the strength of the prediction effects (Hudson et al., 2017). Most recently, we showed that similar effects can be elicited by the prior object context, such that hands reaching straight for an object are perceptually judged to veer slightly upwards if they would need to reach over an obstacle, and slightly downwards when reaching unnecessarily high (Hudson et al., 2018), as long as

intentionality is attributed to the moving stimulus (McDonough, Hudson & Bach, 2019).

These data show that the goals attributed to others are indeed translated into predictions about their upcoming action, which then bias perceptual judgments towards these expectations. Yet, in all these studies the goals or environmental constraints were explicitly given prior to action onset. In the real world, people typically do not always announce their intentions before action. Instead, an action's goal often has to be derived dynamically once the action is underway and its kinematics become apparent (Ambrosini, Costantini, & Sinigaglia, 2011; Ambrosini et al., 2013; Bach, Bayliss, & Tipper, 2011; Sartori, Becchio, & Castiello, 2011). We have argued that the affordances of the goal objects could play a major role in deriving a person's goals once an action is underway (Bach et al., 2014). Viewing an action (e.g. a hammering motion) that matches the affordances of an available goal object (e.g. a hammer) would immediately signal to an observer what the goal of the action would be (Bach et al., 2005; van Elk, van Schie, & Bekkering, 2014). And indeed, there is now ample evidence that people spontaneously derive the target of a reach, by matching the hand's grip configuration – i.e. either a small "precision" grip or a large "power" grip – to the available large or small objects in the environment (Ambrosini et al., 2011; Ambrosini et al., 2013; for a review, see Bach et al., 2014). For example, eye movements reveal that people anticipate the target of an ongoing reach by matching the unfolding grip shape (large or small grip) to the surrounding objects (Ambrosini et al., 2011; Ambrosini et al., 2013), automatic imitation effects are larger for actions that fit a goal object (Bach et al., 2011) and larger motor evoked potential are elicited by the same kinematics if they fit an available goal object (Southgate, Johnson, Karoui, & Csibra, 2010).

Here we test, for the first time, whether people use such grip-object matching not only to derive the action's goal or target (e.g., which goal object is selected), but whether they also use these goals, as assumed by perceptual prediction models (Kilner et al., 2007ab), to predict how an action is assumed to develop, even if it is already well underway. If this is the case, we should find that the match of an unfolding hand grip to one of two objects in the environment should again induce such perceptual biases, and they should be measureable – as in our prior work – in subtle distortions in perceptual judgment about these actions. Demonstrating such distortions is crucial to show that, during action observation, people go beyond simple goal inference (e.g. identifying the target of a reach) but that they use this information to then predict which future course the action will take, predicting the subtle kinematic change towards the identified goal object.

In two experiments, we presented participants with brief videos of an actor's hand starting at rest and then starting to reach towards the centre point between two adjacent potential target objects – one small, one large –, with the hand forming either a whole-hand power grip or a precision grip. The hand disappeared mid-motion, at an equal distance away from either object, and participants were required to indicate the final location of the hand's index finger on a touch screen monitor. If observers identify the goals of the action by matching the observed grip to the two objects' affordances and then form a perceptual prediction about its future course, then perceptual judgments should show specific biases: the located disappearance points should be reported closer to the corresponding object than they actually were, and away from the alternative (mismatching) target object. Therefore, although the hand actually reached between the two objects, reaches with a precision grip should

be reported closer to the smaller object and reaches with a power grip should be perceptually biased towards the larger object.

A crucial question is whether any such effects emerge from a general top-down mechanism, such that high-level attributions of others' goals penetrate lower-level perceptual representations, or whether any perceptual biases emerge from "encapsulated" interactions in the perceptual system itself (Firestone & Scholl, 2016; Scholl & Gao, 2013), which has already been shown to be sensitive to such matching hand-object interactions (Bracci & Peelen, 2013). We assumed that the perceptual biases in our prior research (Hudson et al., 2016[ab], 2018) emerged from high-level information, when people either heard the person make goal statements ("I'll take that!") or instructed them about the appropriate action ("Take it!"). However, in these studies, these goals were given well in advance of the action commencing, so that participants had ample time to "tune" lower-level processes towards the expected goal. To test whether such high-level information can penetrate online social perception – when the action's goal only becomes apparent while the action is underway – we manipulated, across the two experiments, whether participants had to explicitly derive the action's goals or were given no such instruction. In Experiment 4a, we had participants say into the microphone, after each trial, which object they thought the hand was reaching for, therefore requiring explicit goal monitoring and attribution in each trial. In Experiment 4b, no such verbal responses had to be given but participants still reported the perceived disappearance points. These localisation judgments therefore measure spontaneous and implicit goal inferences and the resulting predictions. The difference between the experiments will reveal the extent to which perceptual biases emerge from a more or less encapsulated, automatic visual predictions system that relies on perceptually available "local" stimulus

features (e.g., hand pre-shape, available objects, for a review, see Scholl & Gao, 2013), or whether these processes can be penetrated by higher-level information, such as the explicit attribution of goals to another. Moreover, the combination of explicit verbal goal judgments and implicit perceptual judgments in Experiment 4a will also allow us to test, across participants and across trials, the relationship between these measures.

6.2.1 Method

6.2.1.1 Participants

Sixty-two participants took part in Experiment 4a (mean age = 20 years, SD = 3.4, 52 females) and 63 participants took part in Experiment 4b (mean age = 21 years, SD = 5.5, 50 females). Eleven additional participants across both experiments were excluded due to performance assessed against several a priori criteria (see Results). All were right handed and had normal/corrected-to-normal vision, and were recruited from Plymouth University for course credit. The study was approved by the University of Plymouth Ethics Committee, in accordance with the declaration of Helsinki. A power analysis revealed that a sample size of 62 provides .80 power to detect effects in the predicted direction with Cohen's *d* = .31, and effects in either direction with Cohen's d = .36. Our prior studies investigating similar effects with the same method (Hudson et al., 2017; Hudson et al., 2018; McDonough et al., 2019) revealed that effect sizes are consistently larger (*d* = .52 to *d* = 1.23).

6.2.1.2 Apparatus

Presentation (NeuroBS) software was used to present the experiment via a HP EliteDisplay S230tm 23-inch widescreen (1920 x 1080) Touch Monitor. Verbal responses for Experiment 4a were detected using Presentation's Sphinx speech recognition engine via a Microsoft LifeChat LX-3000 Headset.

6.2.1.3 Stimuli

Example stimuli can be seen in Figure 6-1A. Stimuli were derived, using photo manipulation, from a prior stimulus set of video stimuli from one of the authors (Costantini, Ambrosini, & Sinigaglia, 2012). The videos (950x540) showed an actor's arm, from the side view, reaching towards a location in-between a small target (a strawberry) and a large target (an apple). They were derived from videos of natural (e.g. non-pantomimed) reaches towards one of the two objects, which were then photo-edited such that both objects were located at an equal distance away from the hand, with one closer to the foreground and lower down on the screen and one closer to the background and higher up on the screen (object positions counterbalanced across trials). The two objects were chosen because both their size and typical use clearly affords small "precision" grips (strawberry) and large whole-hand "power" grips (the apple), but they have similar abstract round shapes and merely differ in colour, so that they represent prototypical objects.

The actor's hand started at rest in a neutral closed hand posture, and then began to reach, with the hand progressively opening to form either a whole-hand "power" grip or a thumb-index "precision" grip. Four reach videos were used for each hand pre-shape condition, which together with two target layouts (small object to the front,

large at the back, or vice versa), created a total of 16 different videos. Each video was converted into 9 frames, where frame 9 showed the actor's hand at maximal pre-shape, halfway between the starting position and the objects, and halfway between the two objects, so that only pre-shape information would predict which object would be reached for and not the hand position. The shadow of the hand was digitally removed so this information could not aid localisation.

Response images for both experiments were created by digitally removing the actor's arm and hand from the scene, so that only the target objects and the background remained. Presenting this frame immediately after the action sequence gave the impression of the hand disappearing from the scene. A second response image for Experiment 4a was identical to these images, with the addition of four question marks positioned at each corner of the screen. These served as cues for participants to make their verbal responses about which object they believed was the target. All editing was completed using Adobe CC Photoshop.

6.2.1.4 Procedure

An example trial sequence can be seen in Figure 6-1B. Participants completed a total of 192 trials, consisting of four blocks of 48 trials (each representing all 16 different trials three times), with breaks in between. At the start of each trial, participants saw an instruction to "Hold the spacebar", to which they pressed the spacebar with their right hand and kept it depressed until the end of the action sequence. This ensured that they did not track the observed action with their finger, and could only initiate their response once the action sequence had disappeared. They then saw the first (neutral) frame of the action sequence for 1000ms, followed

by successive frames at 80ms intervals. The final frame was randomly chosen in each trial as either frame 8 or frame 9, to increase variability of the hand's final position. This final frame was then immediately replaced with the response image. Participants released the spacebar and, with their right hand, touched the screen where they thought the final position of the tip of the observed index finger was. For Experiment 4a, the touch response was immediately followed by the second response frame where participants were required to say into the microphone which target object they thought the actor was reaching towards (either "apple" or "strawberry"). Once the verbal response was registered, the next trial began. For Experiment 4b, the next trial began as soon as the touch response was registered.

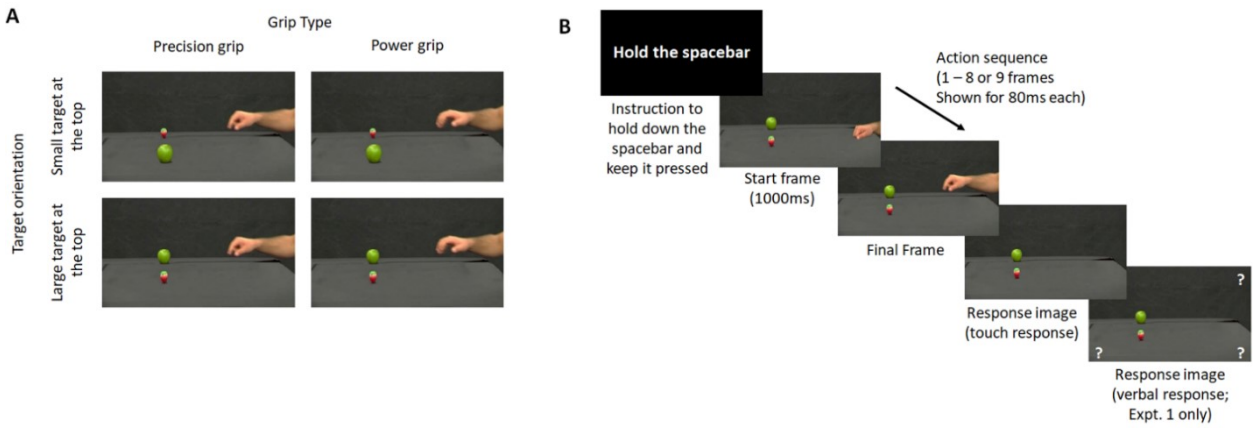


Figure 6-1. Experiments 4a and 4b. Experimental conditions and trial sequence.

Panel A. Experimental conditions: The objects were arranged with either the small object (strawberry) on top and the large object (apple) on the bottom (top row), or with the large object on top and the small object on the bottom (bottom row). The actor's hand reached with either a precision (small) grip (left column), or with a power (large) grip (right column). Panel B. Example of a trial sequence, showing a "large object on top" configuration with a small grip.

6.2.2 Results

Data filtering was identical to Hudson et al. (2018) and McDonough et al. (2019). In both experiments, individual trials were excluded if the correct response procedure was not followed (e.g. lifting the spacebar before the response image as presented; 2.8% of total trials), or if response initiation or execution times were less than 200ms or more than 3SDs above the sample mean (2.4%, Initiation: mean =355.5ms, *SD*=143.7; Execution: mean =646.0ms, *SD*=240.0). Participants were excluded if too few trials remained after trial exclusions (<50% trials, 5 participants), if their average distance between the real and selected positions was more than 3SDs away from the sample mean (mean =39.0 pixels, *SD*=17.0, 2 participants excluded), or if the correlation between the real and selected positions was more than 3SD below the median r value (X axis: median *r* =.762, *SD* = .113; Y axis: median *r* =.860, *SD* = .098, 2 participants excluded). Two further participants were excluded from Experiment 4a, one because they selected the top object as the most likely target object in *all* trials, and one for showing an abnormally large effect in the predicted direction (e.g. 15 times larger than the sample mean) so that we suspected a misunderstanding of the task (e.g. touching the likely target object instead of the hand disappearance point). Removal of these two participants does not affect the results. This left a total of 62 participants in Experiment 4a and 63 participants in Experiment 4b.

Analysis was conducted on the perceptual bias, which reflects the difference between the hand's real disappearance point and participants' subjective judgments. It was derived by subtracting the real final coordinates of the tip of the index finger from the participant's selected coordinates on each trial (see Figure 6-2). This

resulted in separate difference scores for the X and Y axis, where positive X and Y scores represented a rightward and upward displacement respectively, and negative X and Y scores represented a leftward and downward displacement respectively. A score of 0 on both axes indicated that the participant selected the real final position exactly.

6.2.2.1 Y axis

Participants' perceptual biases on the Y axis were analysed with a 2x2x2 mixed ANOVA, with Grip type (power vs precision) and Object location (large target on top vs small target on top) as repeated measures factors and Experiment (4a: explicit prediction vs 4b: implicit prediction) as between-subjects factor. We predicted, first, that perceived disappearance points would be distorted towards the apparent target object of the reach, such that reaches would appear to have terminated slightly higher if they match a target object at the top and lower for a target object at the bottom. Second, if these perceptual biases are informed by – or do inform – explicit judgments, then these shifts should be larger, or only observed, in Experiment 4a, where the hands' goals were task relevant, compared to Experiment 4b where such goal inferences would need to be made spontaneously, in parallel to the perceptual judgment task. These should be reflected in a two-way interaction of Grip type and Object location and a three-way interaction of Grip type, Object location, and Experiment.

The analysis revealed the predicted interaction of Grip type and Object location, $F(1,123) = 16.8$, $p < .001$, $\eta_p^2 = .120$. Even though kinematics were identical, the disappearance point of power grips was reported higher when the large target object

was placed at the top (-15.1px) than when the large target object was placed at the bottom (-16.0px, $t(124)=2.75$, $p=.007$, $d=.25$). Conversely, the disappearance point for precision grips was reported to be higher when the small target object was at the top (-1.5px) compared to when the small target object was at the bottom (-2.8px, $t(124)=3.64$, $p<.001$, $d=.33$). Note that while these deviations are small, they are highly reliable, reflected in a Bayes factor of BF10 of 130.5 for the interaction, and – especially when their smaller than lifelike size is considered – surpass similar deviations induced by object biases in the kinematics of real reaches (e.g., distractor interference/deviation ≈ 1mm, (Keulen, Adam, Fischer, Kuipers & Jolles, 2003; Welsh, Elliott & Weeks, 1999).

The second question was whether the size of these perceptual displacements was larger when the actions' target was task relevant in Experiment 4a compared to when it was task-irrelevant in Experiment 4b. However, there was no three-way interaction between grip type, object location and experiment ($F(1,123) = .666$, $p = .416$, $\eta_p^2 = .005$). Moreover, the relevant interaction of grip and object location was present in both experiments, irrespective of whether participants explicitly reported the action's goals after the perceptual judgments (Experiment 4a: $F(1,61) = 9.56$, $p = .003$, $\eta_p^2 = .135$, Experiment 4b: $F(1,62) = 7.20$, $p = .009$, $\eta_p^2 = .104$, see Figure 6-2). Two one-sided tests (TOST) procedure (Lakens, 2017) indicated that the observed effect size of the between-experiment difference ($d=.16$) was significantly within the equivalence bounds of $\Delta L = –.51$ and $\Delta U = .51$, $t(113.89) = -1.95$, $p = .027$, and a Baysian analysis provides substantial evidence for the null hypothesis, BF10 = .110 , There were no further main effects or interactions (all $F < 1.62$, all $p > .205$).

As all other effects in the ANOVA were not predicted, they should be treated as incidental findings, unless they pass a threshold of p < .01, corrected for multiple comparisons in an ANOVA (Cramer et al., 2015). Only one, a main effect of grip type, $F(1,123) = 1130$, $p < .001$, $\eta_p^2 = .902$, surpassed this threshold, showing perceived disappearance points of power grips were displaced further downward than precision grips. This was expected since the power grip is larger and therefore has a lower centre of gravity, which is known to affect touch screen judgements, but is independent from our effects of interest (Coren & Hoenig, 1972; see also Hudson et al., 2017; Hudson et al., 2018).

6.2.2.2 X axis

Perceptual biases on the X axis were analysed with the same ANOVA model. As we did not have any prediction for the X axis, all effects are therefore subject to alpha inflation due to multiple comparisons in an ANOVA (Cramer et al., 2015) and should be evaluated against a Bonferroni-adjusted alpha of p < .007. Only the main effect of grip type, $F(1,123) = 503$, $p < .001$, $\eta_p^2 = .804$, passed this adjusted threshold, with the perceived disappearance point of power grips more leftward than precision grips, which again reflects leftward centre of gravity for power grips. There were no further main effects or interactions (all $F < 4.07$, all $p > .046$).

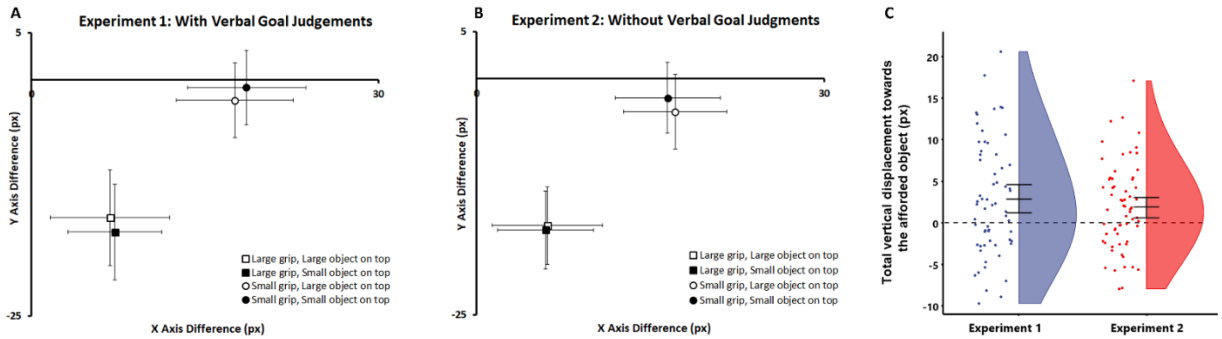


Figure 6-2. Experiments 4a and 4b. Results.

Panel A. Grip type x object interaction for Experiment 4a. The difference scores between the real final position and the selected final position is plotted for the X axis and Y axis. Panel B. Grip type x object interaction for Experiment 4b. Panel C. A raincloud plot (Allen, Poggiali, Whitaker, Marshall & Kievit, 2018) of the comparison across experiments of the size of the Y axis interaction in pixels, equivalent to the total amount by which each grip type was distorted towards the congruent object. Each data point represents this Y axis interaction value for each participant. Error bars depict 95% confidence intervals.

6.2.2.3 Relationships between perceptual shifts and explicit goal judgments

An important question is to what extent the perceptual displacements measured above are informed by people's higher-level, explicit judgments about the observed actions. If perceptual predictions are shaped by high-level goal attributions or vice versa, then perceptual displacement and explicit judgments in Experiment 4a should be closely linked, both across participants and across trials within participants.

6.2.2.3.1 Do people rely on affordance matching to make explicit goal judgments?

We first established whether grip/object-matching does not only inform perceptual biases in action observation, but also people's *explicit* judgments about the action's goals. Participants were never given explicit instruction about the relevance of grip/object match, but they reported, after each action in Experiment 4a, whether

they subjectively experienced the hand to be reaching for the apple or the strawberry. To test whether these judgments were informed by the apparent grip/object-match, we separately coded which object the grip *actually* corresponded to and participants' subjective judgments about which object they *felt* the hand reached for. To this end, a hand grip matching the bottom object was coded as 0 and a match to the top object was coded as 1. Verbal goal judgments were similarly coded as 0 and 1 for perceived goal objects at the bottom and at the top, respectively. We then simply, for each participant, calculated the proportion of verbal goal judgments that corresponded to the actual match with the goal object. A simple t-test against chance (50%) revealed that explicit judgments corresponded well with the actual hand-object match (M = 66.7%, SD = 16.7%; t =31.5, p <.001). This confirms that the grip-object-match did not only inform perceptual displacements (see main analysis), but also participants' explicit goal object judgments.

6.2.2.3.2 Are perceptual biases and explicit judgments related across participants?

We then tested whether individual differences in participants' tendency to rely on grip/object matching to make explicit judgments is related to their reliance on grip/object matching in perceptual judgments. We therefore correlated the proportion to which each participant's verbal goal judgments matched the actual hand/object match with their perceptual biases towards the grip-matching goal object (i.e. the interaction contrast that marks the predictive perceptual shift due to matching grips to object affordances in the main analysis above). Surprisingly, the two types of judgments were almost perfectly uncorrelated, $r(59) = .08$, $p = .518$, N=62. TOST procedure (Lakens, 2017) indicated that the observed effect size (r=.08) was

significantly within the equivalence bounds of ΔL = –.36 and ΔU =.36, p = .011. A Bayesian analysis provided “substantial” evidence for the absence of a correlation, BF10 = .102.

6.2.2.3.3 Are perceptual biases and explicit judgments related across trials?

While there may be no overall relationships between a participants’ perceptual biases and explicit judgment, it is possible that such relationships are present on a trial-by-trial basis. If explicit and perceptual judgments depend on one another, then actions/trials judged explicitly to be directed towards the top object should also show a perceptual mis-location towards the top, and vice versa for reaches judged to be directed to the bottom object.

To test this, we first tested whether the perceptual shifts in each trial reflected the *actual* target object location. We therefore correlated, for each participant separately, the actual target object location (coded as 0 or 1) for each trial with the size of the perceptual judgment displacement on the Y axis across all trials of the participant. Testing the resulting fisher-transformed correlation coefficients against zero with a simple t-test, revealed a positive mean correlation between perceptual shifts and target object location across participants (mean r = .03, t = 3.10, p = .003; d = .40), see Figure 6-3A. This trial-by-trial correlation was replicated in Experiment 4b (mean r = .02, t = 2.77; p = .007; d = .35), and did not differ from Experiment 4a (t(123)=.512, p=.609, d=.14). TOST procedure (Lakens, 2017) indicated that the observed effect size (d=.14) was significantly within the equivalence bounds of ΔL = –.51 and ΔU =.51, t(121.5) = -2.06, p = .021. As before, a Bayesian analysis of this effect provided substantial evidence for an absence of such a relationship, BF10 =

.134. Replicating the results of the main analyses with an across-trials correlational measure, this analysis therefore confirms that actions in which the hand grip matched the object on the top, compared to a grip match to the bottom object, induced larger shifts upwards.

Next, we performed an identical across-trials correlation analysis for the relationship between the actual target object location and verbal goal judgment. This again revealed a positive correlation between verbal goal judgments and grip information (mean r = .44, t = 7.20, p <.001; d = .91), see Figure 6-3B. Across trials, reaches whose grip matched the top object were therefore more likely to be judged to be reaching to the top object, and vice versa for reaches whose grip matches the bottom object.

Finally, the crucial question was whether explicit verbal goal judgments about an action and the perceptual shifts showed a positive relationship. Strikingly, as in the across-participants analysis, there was no correlation between perceptual displacements and the explicit verbal goal judgments across trials (mean r = .01, t = .694, p = .500, d = .09), see Figure 6-3C. TOST procedure (Lakens, 2017) indicated that the observed effect size (d=.09) was significantly within the equivalence bounds of ΔL = –.36 and ΔU =.36, t(61) = -2.15, p = .018, and a Bayesian analysis provided strong evidence for the absence of a relationship, BF10 = .088. Thus, while the actual target location informed perceptual judgments and verbal goal judgments, the two types of judgments were not related to each other.

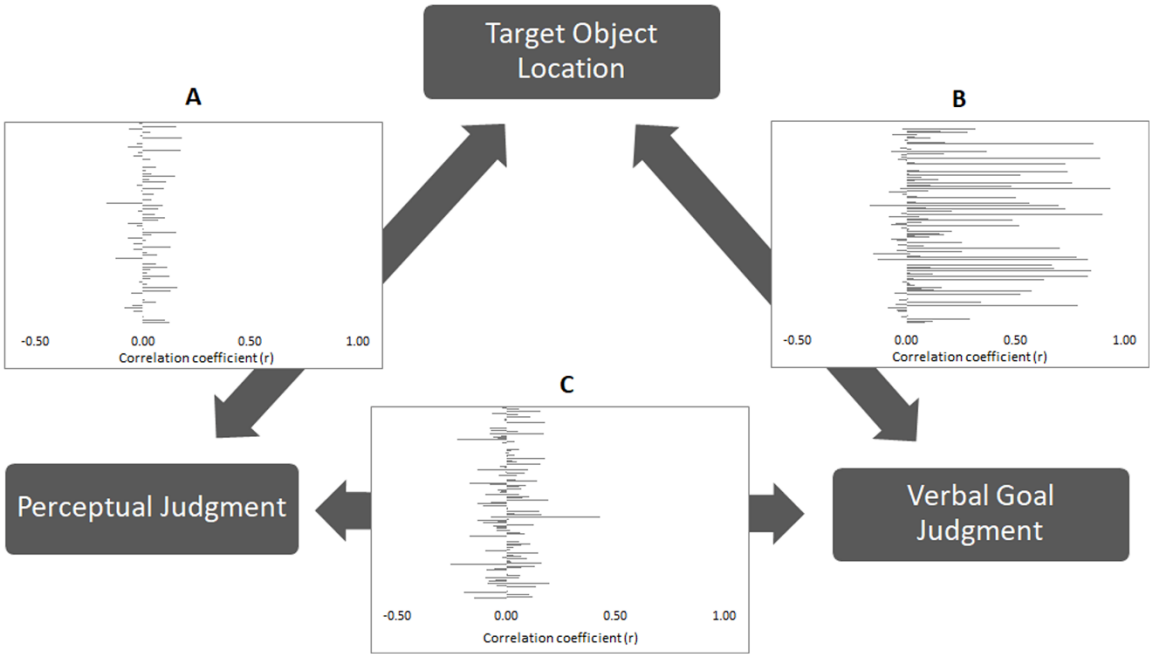


Figure 6-3. Experiments 4a and 4b. Correlational results.

Figure 3. Each participant's across-trial correlation coefficient (r) for the correlation between perceptual judgments and the target object location (top object or bottom object, based on the grip-object match, Panel A), the correlation between verbal goal judgments and the target object location (Panel B), and the correlation between perceptual judgments and verbal goal judgments (Panel C).

6.2.3 Discussion

Prior work has shown that people integrate object and action kinematic information to derive the likely goal of observed actions, even while the action is still ongoing (Ambrosini et al., 2011; Bach et al., 2011; Decroix & Kalénine, 2018; Eshuis, Coventry, & Vulchanova, 2009; for a review, see Bach et al., 2014). Here, we tested the hypothesis of predictive processing models (Bach & Schenke, 2017; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016), that such goal inferences are immediately translated into perceptual predictions about the actions future path towards the inferred goal, and bias perceptual judgments towards these expected trajectories.

The data from two experiments supported this proposal. In each trial, participants observed the initial stages of a reach towards two potential target objects that differed in size, with either a whole-hand power grip or a precision grip, and were asked to perceptually report the hand's last seen position after its sudden offset. The results revealed consistent biases in perceptual judgments towards action expectations derived from the compatibility between the emerging grip type and object size. While reaches with a power grip were reported to be closer to large objects, reaches with a precision grip were perceived to be closer to small objects, even when actions with the same kinematics were observed and only the location of the relevant target object changed. These perceptual mis-locations were present both when participants were explicitly asked to identify the goal objects in a secondary task (Experiment 4a) and when the reach targets were completely task irrelevant and participants were only asked to accurately report the hand's disappearance point (Experiment 4b). Moreover, they were observed even though the actor's hand started at rest before pre-shaping. The target object was therefore ambiguous before action onset and only became apparent once the action commenced and a specific grip type began to form.

These perceptual displacements towards the expected kinematics support predictive processing models of social perception, which argue that any inferences about an observed action's goal will (1) give rise to predictions about the action's further kinematics, which can then (2) bias action perception towards these expectations (Bach & Schenke, 2017; Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016; Kilner et al., 2007a, 2007b). They go beyond previous findings in which action expectations were explicitly induced prior to action onset, for example, by asking participants to instruct the (virtual) actor

(Hudson, Nicholson, Ellis, et al., 2016), by hearing goal statements of the actor (“I’ll take it!”, Hudson et al., 2017; Hudson, Nicholson, Simpson, et al., 2016), or by presenting a static image of the goal object with or without obstructing objects in the way (Hudson et al., 2018; McDonough et al., 2019). Here, no such prior information was available. The actions started from a neutral position and the goals only became apparent once it was underway, from the subtle pre-shaping of the hands for the affordances of the goal object (i.e. precision grip when directed towards the small object, power grip when directed towards the large object).

Our results therefore show, first, that predictions are not just made before action onset, but are dynamically adjusted “on-line” as more information becomes available from the unfolding kinematics and are then integrated with the action’s perceptual representation. Second, they reveal that matching of actions to the affordances of potential goal objects in the environment plays a major role in this process, as previously hypothesized (see Bach et al., 2014 for a theoretical proposal and review). Third, they go beyond prior work that has shown that people use such affordance matching to identify the *target* of another’s action, guiding eye movements towards it in an anticipatory manner (Ambrosini et al., 2011; Ambrosini et al., 2013; Bach et al., 2011). Instead, they reveal that these predictions do not just represent the likely goal object, but represent concrete expectations about the next step of a hand’s path through the scene, which interact with the perceptual representation of the kinematics that were actually observed.

A surprising finding was that this matching of actions to goal objects and the resulting perceptual biases appeared to be highly automatic and independent from explicit judgments. We had hypothesized that if perceptual mislocations and explicit

judgments inform each other, then those trials that were explicitly judged to be directed towards the top object should also show upwards mis-locations, and vice versa for judgments towards bottom objects. However, converging findings from three independent tests argue against this interpretation. First, in Experiment 4a, the action's goal was highly task relevant because participants indicated verbally, after each action, which object they believed was the target. No such response was required in Experiment 4b, such that any perceptual bias therefore indexes only spontaneous, implicit goal inferences and prediction. Nevertheless, the perceptual bias towards the grip-matching object was evident – with virtually identical effect sizes – in both experiments. Second, in Experiment 4a, correlational analyses across participants showed that the perceptual biases were independent of whether participants' verbal goal judgments revealed a reliance on grip information or not. Thus, the perceptual bias towards the matching goal object was of similar size irrespective of whether participants made use of grip-object matching when they explicitly judged the action's goals. Third and finally, correlational analyses across trials that directly relate perceptual displacements in a given to trial to verbal goal judgments in the same trial (Experiment 4a) confirmed this lack of a top-down influence. Even though the actual goal object predicted both the direction of the perceptual bias and which object was explicitly reported as a target, the perceptual biases and verbal goal judgments remained uncorrelated across trials. In other words, while the mechanisms for action prediction and the goal identification both rely on grip-object matching, the two mechanisms do not strongly inform each other: explicit goal judgments do not induce perceptual biases, nor do perceptual biases induce explicit goal judgments. While it had been difficult to draw conclusions from null effects, recent Bayesian and equivalence-testing (TOST) analysis techniques

have overcome this problem (Lakens, 2017). Indeed, all three tests described above provide substantial to strong evidence *for* the absence of a link between explicit judgments and perceptual biases, and the reported difference remain significantly within equivalence bounds.

This apparent dissociation between explicit and implicit perceptual biases may appear surprising from a viewpoint of predictive coding models, according to which predictions and prediction errors ensure that top-level and lower-level judgments remain aligned (Clark, 2013; Friston & Kiebel, 2009; Hudson et al., 2017). Thus, any inferences on a higher level – for example, what the goal of the action is – would propagate downwards to lower levels and inform perceptual judgments. Conversely, any change in perceptual estimation – whether it is perceived to travel upwards and downwards – would, via prediction errors, inform resulting high-level judgments of action goals. This apparent conflict can be resolved, however, if one accepts recent proposals that predictions can also emerge locally, from top-down interactions *within* the human perceptual system for the perception of biological motion (for a review, see Scholl & Gao, 2013), without drawing on information external to these networks such as high-level explicit action goal judgments (Firestone & Scholl, 2016). Indeed, several lines of evidence suggest that the perceptual system itself can detect many aspects of intentional behaviour, without the need for higher-level evaluation, such as whether one actor chases another (Gao, McCarthy & Scholl, 2010), whether an actor pays attention to their reach or whether an actor moves certain limbs in a particular direction, dependent on this attention orientation (e.g. moving an arm to the left when the actor is attending to this target position, Jellema, Baker, Wicker, & Perrett, 2000). The match between hand and goal object may therefore provide another feature from which such lower-level teleological interpretations of observed

motion can be derived, and low-level perceptual regions have indeed been found to be sensitive to such information (e.g., STS, Gao, Scholl & McCarthy, 2012; Saxe, Xiao, Kovacs, Perrett & Kanwisher, 2004; lateral occipital cortex, Bracci & Peelen, 2013). Our new data then suggests that these regions are not only sensitive to the presence of these matches and the goals they signify, but that they also use them to predict the action's further path and bias the perceptual representations towards it, independently of the goals *explicitly* attributed to the other person.

Future studies now need to investigate from what kind of mechanism the perceptual biases emerge. Our prior studies point towards lower-level perceptual processes that determine participants' conscious perceptual experience of the actions, which then drives their explicit judgments. First, in our original studies (Hudson et al., 2017; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016), these effects were measured not with touch screen judgments, but probe judgments. Participants compared the hand disappearance point with probes in the same, forward, or prior position. Even when these probes were presented only 250 ms after hand disappearance, the perceptual distortions were apparent, suggesting, at the very least, an effect in iconic memory. Second, in our most recent work testing perceived changes to action kinematics in the presence of obstacles, all such effects were eliminated when dynamic visual noise masks were presented briefly (560 ms) after action offset (Hudson et al., 2018), which are known to interfere with re-entrant top-down projections to early visual cortex (Fahrenfort, Scholte, & Lamme, 2007; Lamme, Zipser, & Spekreijse, 2002). Together, therefore, these findings support a low-level locus of the effects that either reflects the top-down sharpening of the uncertainty during motion perception (i.e. motion blurring, Hammett, 1997), or the filling-in of the expected path after the unexpected sudden offset (Ekman et al.,

2017). Neuroimaging studies would be useful to disentangle to what extent the perceptual changes we have measured here reflect changes to early perceptual systems, similar to that seen in various visual illusions and sometimes in motion illusions (e.g. apparent motions, Muckli, Kohler, Kriegeskorte, & Singer, 2005; predicted motion pre-play, Ekman et al., 2017).

6.2.4 Conclusions

The present results reveal that the perceptual experience of others' actions is predictively shaped by the integration of the unfolding action kinematics with the affordances of available goal objects, as proposed by recent predictive models of social perception, (Csibra, 2008; Kilner et al., 2007[ab]; Bach, Nicholson & Hudson, 2014; Schenke & Bach, 2017). These integrations likely emerge at a relatively low-level, from processes within systems for the perception of biological motion, without influences from top-down evaluations of others' goals and intentions. Future studies must now resolve precisely via which mechanism predictions act on perceptual representations, how they help guide own actions towards future states in social interactions, and how prior knowledge is updated and revised if it consistently fails to explain the perceptual input.

6.2.5 References

Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion, 30*(2), 177-186. doi:10.1007/s11031-006-9020-2

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7*(10), 1057-1058. doi:10.1038/nn1312

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. (2018). Raincloud plots: a multi-platform tool for robust data visualization. *PeerJ Preprints, 6,* e27137v1.

Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neurophysiology, 106*(3), 1437-1442. doi:10.1152/jn.00118.2011

Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking Ahead: Anticipatory Gaze and Motor Ability in Infancy. *PLOS ONE, 8*(7), e67916. doi:10.1371/journal.pone.0067916

Ansuini, C., Cavallo, A., Bertone, C., & Becchio, C. (2015). Intentions in the brain: the unveiling of Mister Hyde. *The Neuroscientist, 21*(2), 126-135.

Bach, P., Bayliss, A. P., & Tipper, S. P. (2011). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychonomic Bulletin & Review, 18*(1), 171-176. doi:10.3758/s13423-010-0029-x

Bach, P., Knoblich, G., Gunter, T. C., Friederici, A. D., & Prinz, W. (2005). Action Comprehension: Deriving Spatial and Functional Relations. *Journal of Experimental Psychology: Human Perception and Performance, 31*(3), 465-479. doi:10.1037/0096-1523.31.3.465

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Frontiers in Human Neuroscience, 8*. doi:10.3389/fnhum.2014.00254

Bach, P., & Schenke, K. C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass, 11*(7), e12312. doi:10.1111/spc3.12312

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617-629. doi:10.1038/nrn1476

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature, 402*(6764), 877-879. doi:10.1038/47245

Bracci, S., & Peelen, M. V. (2013). Body and Object Effectors: The Organization of Object Representations in High-Level Visual Cortex Reflects Body-Object Interactions. *Journal of Neuroscience, 33*(46), 18247-18258. doi:10.1523/jneurosci.1322-13.2013

Chetverikov, A., & Ivanchei, I. (2016). Seeing “the Dress” in the Right Light: Perceived Colors and Inferred Light Sources. *Perception, 45*(8), 910-930. doi:10.1177/0301006616643664

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(03), 181-204. doi:10.1017/s0140525x12000477

Coren, S., & Hoenig, P. (1972). Effect of Non-Target Stimuli upon Length of Voluntary Saccades. *Perceptual and Motor Skills, 34*(2), 499-508. doi:10.2466/pms.1972.34.2.499

Costantini, M., Ambrosini, E., & Sinigaglia, C. (2012). Out of your hand's reach, out of my eyes' reach. *Quarterly Journal of Experimental Psychology, 65*(5), 848-855. doi:10.1080/17470218.2012.679945

Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., . . . Wagenmakers, E.-J. (2015). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review, 23*(2), 640-647. doi:10.3758/s13423-015-0913-5

Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorymotor foundations of higher cognition. Attention and performance XXII* (pp. 435-459). Oxford: Oxford University Press.

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences, 22*(9), 764-779. doi:10.1016/j.tics.2018.06.002

Decroix, J., & Kalénine, S. (2018). Timing of grip and goal activation during action perception: a priming study. *Experimental Brain Research, 236*(8), 2411–2426. doi:10.1007/s00221-018-5309-0

Donnarumma, F., Costantini, M., Ambrosini, E., Friston, K., & Pezzulo, G. (2017). Action perception as hypothesis testing. *Cortex, 89*, 45-60. doi:10.1016/j.cortex.2017.01.016

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications, 8*, 15276. doi:10.1038/ncomms15276

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009). Predictive Eye Movements Are Driven by Goals, Not by the Mirror Neuron System. *Psychological Science, 20*(4), 438-440. doi:10.1111/j.1467-9280.2009.02317.x

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2007). Masking Disrupts Reentrant Processing in Human Visual Cortex. *Journal of Cognitive Neuroscience, 19*(9), 1488-1497. doi:10.1162/jocn.2007.19.9.1488

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences, 39*. doi:10.1017/s0140525x15000965

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(1), 126-132. doi:10.1037//0278-7393.10.1.126

Frischen, A., & Tipper, S. P. (2006). Long-term gaze cueing effects: Evidence for retrieval of prior states of attention from memory. *Visual Cognition, 14*(3), 351-364. doi:10.1080/13506280544000192

Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks, 22*(8), 1093-1104. doi:10.1016/j.neunet.2009.07.023

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493-501.

Gao, T., McCarthy, G., & Scholl, B. J. (2010). The Wolfpack Effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science, 21*(12), 1845-1853. doi:10.1177/0956797610388814

Gao, T., Scholl, B. J., & McCarthy, G. (2012). Dissociating the Detection of Intentionality from Animacy in the Right Posterior Superior Temporal Sulcus. *Journal of Neuroscience, 32*(41), 14276-14280. doi:10.1523/jneurosci.0562-12.2012

Hammett, S. T. (1997). Motion blur and motion sharpening in the human visual system. *Vision Research, 37*(18), 2505-2510. doi:10.1016/s0042-6989(97)00059-x

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12*(5), 822-851. doi:10.3758/bf03196775

Hudson, M., Bach, P., & Nicholson, T. (2017). You said you would! The predictability of other's behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance, 44*(2), 320-335. doi:10.1037/xhp0000451

Hudson, M., McDonough, K. L., Edwards, R., & Bach, P. (2018). Perceptual Teleology: Expectations of Action Efficiency Bias Social Perception.

Proceedings of the Royal Society: B.(285(1884)), 20180638. doi:10.1098/rspb.2018.0638

Hudson, M., Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition, 146*, 245-250. doi:10.1016/j.cognition.2015.09.021

Hudson, M., Nicholson, T., Simpson, W. A., Ellis, R., & Bach, P. (2016). One step ahead: The perceived kinematics of others' actions are biased toward expected goals. *Journal of Experimental Psychology: General, 145*(1), 1-7. doi:10.1037/xge0000126

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences, 9*(1), 21-25.

Jacquet, P. O., Chambon, V., Borghi, A. M., and Tessari, A. (2012). Object affordances tune observers' prior expectations about tool-use behaviors. *PLoS One 7*:e39629. doi: 10.1371/journal.pone.0039629

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural Representation for the Perception of the Intentionality of Actions. *Brain and Cognition, 44*(2), 280-302. doi:10.1006/brcg.2000.1231

Joyce, K., Schenke, K., Bayliss, A., & Bach, P. (2015). Looking ahead: Anticipatory cueing of attention to objects others will look at. *Cognitive Neuroscience, 7*(1-4), 74-81. doi:10.1080/17588928.2015.1053443

Kalénine, S., Shapiro, A. D., Flumini, A., Borghi, A. M., & Buxbaum, L. J. (2014). Visual context modulates potentiation of grasp types during semantic object categorization. *Psychonomic bulletin & review,* 21(3), 645-651.

Keulen, R. F., Adam, J. J., Fischer, M. H., Kuipers, H., & Jolles, J. (2003). Distractor interference in selective reaching: Dissociating distance and grouping effects. *Journal of motor behavior, 35*(2), 119-126.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007a). The mirror-neuron system: a Bayesian perspective. *NeuroReport, 18*(6), 619-623. doi:10.1097/wnr.0b013e3281139ed0

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007b). Predictive coding: an account of the mirror neuron system. *Cognitive Processing, 8*(3), 159-166. doi:10.1007/s10339-007-0170-2

Koster-Hale, J., & Saxe, R. (2013). Theory of mind: a neural prediction problem. *Neuron, 79*(5), 836–48. doi:10.1016/j.neuron.2013.08.020

Lakens, D. (2017). Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355-362. doi: 10.1177/1948550617697177

Lamme, V. A. F., Zipser, K., & Spekreijse, H. (2002). Masking Interrupts Figure-Ground Signals in V1. *Journal of Cognitive Neuroscience, 14*(7), 1044-1053. doi:10.1162/089892902320474490

McDonough, K. L., Hudson, M., & Bach, P. (*accepted*). Cues to intention bias action perception toward the most efficient trajectory. *Scientific Reports*.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary Visual Cortex Activity along the Apparent-Motion Trace Reflects Illusory Perception. *PLoS Biology, 3*(8), e265. doi:10.1371/journal.pbio.0030265

Nicholson, T., Roser, M., & Bach, P. (2017). Understanding the Goals of Everyday Instrumental Actions Is Primarily Linked to Object, Not Motor-Kinematic, Information: Evidence from fMRI. *PLOS ONE, 12*(1), e0169700. doi:10.1371/journal.pone.0169700

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370*(6487), 256-257. doi:10.1038/370256b0

Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences, 31*(02). doi:10.1017/s0140525x08003804

Pellicano, E., & Burr, D. (2012). When the world becomes “too real”: A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences, 16*(10), 504–510. doi:10.1016/j.tics.2012.08.009

Prinz, W., & Rapinett, G. (2008). Filling the gap: Dynamic representation of occluded action. *Emerging Communication, 10*(223).

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. Nat Rev Neuro, 11, 264-274. doi: 10.1038/nrn2805.

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., & Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiological Reviews, 94(2),* 655-706.

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition, 119*(2), 242-252. doi:10.1016/j.cognition.2011.01.014

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional

actions. *Neuropsychologia, 42*(11), 1435-1446. doi:10.1016/j.neuropsychologia.2004.04.015

Schenke, K. C., Wyer, N. A., & Bach, P. (2016). The Things You Do: Internal Models of Others' Expected Behaviour Guide Action Observation. *PLOS ONE, 11*(7), e0158910. doi:10.1371/journal.pone.0158910

Scholl, B. J., & Gao, T. (2013). Perceiving Animacy and Intentionality: Visual processing or higher-level judgment. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social Perception: Detection and interpretation of animacy, agency, and intention,* (pp. 197-230): The MIT Press.

Southgate, V., Johnson, M. H., Karoui, I. E., & Csibra, G. (2010). Motor System Activation Reveals Infants' On-Line Prediction of Others' Goals. *Psychological Science, 21*(3), 355-359. doi:10.1177/0956797610362058

Uithol, S., van Rooij, I., Bekkering, H., & Haselager, P. (2011). What do mirror neurons mirror? *Philosophical Psychology, 24*(5), 607-623.

van Elk, M., van Schie, H., & Bekkering, H. (2014). Action semantics: A unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Physics of Life Reviews, 11*(2), 220-250. doi:10.1016/j.plrev.2013.11.005

Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Experimental Brain Research, 211*(3-4), 517-530. doi:10.1007/s00221-011-2706-z

von der Lühe, T., Manera, V., Barisic, I., Becchio, C., Vogeley, K., & Schilbach, L. (2016). Interpersonal predictive coding, not action perception, is impaired in autism. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1693), 20150373. doi:10.1098/rstb.2015.0373

Welsh, T. N., Elliott, D., & Weeks, D. J. (1999). Hand deviations toward distractors Evidence for response competition. *Experimental Brain Research, 127*(2), 207-212.

7 Chapter Seven - General Discussion

The ability to understand the actions of others and engage in social interaction lies at the core of our existence. Understanding the goals, beliefs and attitudes of other people not only informs our expectations of their next movement, but also allows us to prepare our own action in response, allowing for fluid social interaction (Hamilton & Grafton, 2007; Sebanz & Knoblich, 2009). The aim of this thesis was to investigate the mechanisms that underlie this capacity. While traditional models assume a post-action evaluation of the observed behaviour, deriving action goals through mirrored motor activations (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese & Sinigaglia, 2010; Iacoboni, 2009; Rizzolatti & Craighero, 2004), this thesis investigated more recent claims of predictive social perception. Such predictive models suggest instead that the goals and intentions of other's actions can be inferred by drawing upon prior expectations, derived from knowledge about the world and other people (Bach et al., 2014; Csibra, 2008; Kilner, Friston, & Frith, 2007ab). This knowledge is translated into a prediction of the forthcoming action, providing a reference with which to compare the observed action as it unfolds, confirming prior goal assumptions or signalling the need for revision.

Utilising the well-established Representational Momentum paradigm that reliably measures the predictive influences in motion perception (Freyd & Finke, 1984; Hubbard, 2005, 2015), this thesis provides evidence that observers actively predict the upcoming actions of others in such a manner. They show that action expectations rely on the principle of efficient action, providing a reference signal that guides the perception of others' actions and bias their perceptual representation (Chapter Three). These reference signals are perceptually represented (Chapter

Four) and rely on cues to intention (Chapter Five). The results further demonstrate that action predictions do not only rely on knowledge and contextual information that is available prior to action onset, but that they are also dynamically updated during ongoing action observation (Chapter Six). Together, the findings support a predictive processing account of social perception, akin to those that underlie perception in general, and reveal how we so effortlessly navigate the social world (Bach et al., 2014; Csibra, 2008; Kilner, Friston, & Frith, 2007ab).

7.1 Summary of results

Initial experiments, presented in Chapter Three, tested whether human observers interpret the actions of others under the principle of efficient action, predicting them to take the most energy-efficient path to achieve their goals (Csibra & Gergely, 2013; Dennett, 1987; Gergely & Csibra, 2003). Using a Representational Momentum paradigm (Freyd & Finke, 1984; Hubbard, 2005, 2015), they investigated whether these predictions are represented in a perceptual format, providing a reference image for the observed action that distorts its perceptual representation.

Participants observed videos of an actor reaching towards a target object. The actor either reached efficiently (straight towards a target object or arched over an obstacle) or inefficiently (straight towards the obstacle or arched over an empty space). The hand disappeared mid-way through the action and participants indicating the last seen position of the actor's index finger using a touch screen. Across three experiments, perceptual judgements were consistently biased towards the expected efficient trajectory. Straight reaches towards an obstacle were judged to be higher than efficient straight reaches, in line with the expectation that the actor would have

reached over the obstacle, even though the straight reaches in both conditions were identical. Conversely, arched reaches were reported lower when they reached over an empty space than when they reached efficiently over an obstacle.

These biases were evident when participants passively viewed the actions (Experiment 1a) but increased when participants explicitly acknowledged the context (by stating the presence or absence of obstacles, Experiment 1b), and increased further when they explicitly predicted the most efficient action path (Experiment 1c). These results revealed that human observers, indeed, make predictions about the efficiency of others upcoming actions, and these predictions specifically concern the expected kinematics that a rational actor, who is aware of all environmental constraints, would take. The resulting biases towards the expected kinematics highlight the perceptual nature of these predictions and their use as a reference against which the observed action is compared. Moreover, these predictions were derived spontaneously, and the increase in perceptual biases when predictions were made explicit, prior to action onset, confirms that action efficiency understanding is not only derived from the kinematic components of the observed action, but also from predictions informed by prior knowledge and context.

The experiments in Chapter Four tested whether the biases reported in Experiments 1a-c do indeed reflect changes to the perceptual representation of the action, or whether these effects could be explained by later changes in working memory or the motor components required to make touch-screen judgements (Firestone & Scholl, 2015; Ianì, Mazzoni, & Bucciarelli, 2018). To test this, Experiment 1b was replicated and touch-screen responses were replaced by probe stimuli that appeared directly after action offset (with 250 ms. interval, Experiment 2a) either in the predicted direction (e.g. upwards for inefficient straight reaches) or in the unpredicted direction,

as well as forwards and backwards. Results replicated the main findings, showing that probe stimuli in the predicted direction were consistently misjudged to be in the same position as last seen image of the hand, compared to probes in the unpredicted direction. These findings therefore provide evidence against a working memory or motor-based explanation as prediction effects were apparent immediately after action offset, when working memory does not yet have access to visual representations, and were captured using probe judgements that do not rely on visuospatial motor maps (Kerzel, 2005).

The results from Experiment 2b further confirm that predictions of action efficiency act on early perceptual representation. They tested whether these prediction effects can be disrupted by a dynamic visual noise mask that reliably disrupt such low-level perceptual processes (Breitmeyer & Öğmen, 2006; Fahrenfort, Scholte, & Lamme, 2007). Experiment 1b was again replicated and a short (dynamic checkerboard) mask was inserted into half of the trials, over the last seen position of the hand, directly after action offset. Touch responses in non-masked trials were consistently biased towards efficient action expectations, replicating the original findings, while these perceptual biases were disrupted in masked trials, dramatically reducing the size of the effect. This confirms that the perceptual mask successfully disrupted the predictive processing of action stimuli, confirming the low-level perceptual nature of action predictions.

The experiments in Chapter Five investigated how such predictions of action efficiency emerge and upon what information they rely. Specifically, Experiment 3a, 3b and 3c revealed a reliance upon cues to intentionality. Experiment 3a provided evidence of the robustness of these prediction effects by, first, replicating the original experiment (Experiment 1b) and replicating the resulting perceptual biases. The

magnitude of the perceptual biases were reduced when intentionality cues were removed, first by replacing the actor's hand with a non-agentive ball that moved along the same biological trajectory as the hand (Experiment 3b), and reduced even further by removing this biological motion profile so that the ball now moved along a straight path with a constant speed (Experiment 3c). These findings confirm that predictions of action efficiency do indeed depend on the intentionality of the actor, cued by agency and biological motion, and therefore on the extent to which behaviour is deemed rational and intentional. Action predictions are therefore informed by prior information from knowledge (of agency and efficient action) and context (obstacles present).

The experiments so far investigated predictions of action efficiency whereby all information required to make a prediction were available prior to action onset (i.e. presence of obstacles, agency of stimuli). Experiments in Chapter Six confirmed that action predictions can be generated on-line, once the action has begun, even when goal cues are not available prior to action onset. Participants saw videos of an actor reaching towards two adjacent objects (a small object and a large object) forming either a small precision grip or a large power grip. The hand disappeared as it approached the objects and participants judged the disappearance point on a touch-screen. The results showed that large grips were perceived closer to large objects and small grips were perceived to be closer to small objects, even when the same actions were observed and only the placement of the objects were changed. Surprisingly, the size of these prediction effects did not change when explicit goal judgements were made after each action in Experiment 4a and no relationship was found between perceptual shifts and explicit goal judgements across participants or across trials. These findings reveal that on-line adjustments of predicted actions can

be made based on affordance-matching information revealed during the ongoing action, which biases the perceptual representation of observed actions. These distortions appear to emerge from low-level evaluations of the action kinematics, independent of explicit goal identification.

7.2 Implications for prior research and theorizing

7.1.1 Prediction in social perception

A dominant view of social perception explained action understanding as a purely bottom-up process whereby the goals and intentions of others were realised by directly matching observed behaviours to own motor experience, without feedback or inference (Gazzola & Keysers, 2008; Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010). Despite the seemingly compelling evidence in support of this view, from the discovery of mirror neurons in the monkey brain and apparent homologous systems in humans (Di Pellegrino et al., 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Gazzola & Keysers, 2008; Rizzolatti & Craighero, 2004; Umilta et al., 2001), more recent research has highlighted the involvement of top-down processing in action observation (e.g., Bach, Bayliss, & Tipper, 2011; Liepelt et al., 2008). These findings are consistent with predictive processing models that view perception in general as an iterative process of hypothesis testing and revision (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007ab). Prior knowledge and contextual information is used to make inferences about what is being perceived, which is translated into the sensory information that would be received if these inferences are correct, which can then be compared against – or integrated with – the actual input. In this way, expectations can act as a reference signal that biases perceptual representations of

the input towards expectations. In non-social vision, such predictive models explain, for example, how surrounding illumination is cancelled out in colour perception (Bloj, Kersten, & Hurlbert, 1999), how expectations influence the perception of bi-stable figures (Hohwy, Roepstorff, & Friston, 2008), and how illusory motion arises from the perceptual "filling in" of apparent motion stimuli (Muckli, Kohler, Kriegeskorte, & Singer, 2005). Social perception might be supported by similar predictive mechanisms (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007ab).

Initial studies have used a Representational Momentum paradigm (Freyd & Finke, 1984; Hubbard, 2005, 2015), originally designed to investigate top-down predictive influences on non-social motion perception, to test whether similar predictive effects emerge using social stimuli (Hudson, Bach, & Nicholson, 2018; Hudson & Jellema, 2011; Hudson, Liu, & Jellema, 2009; Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016). Hudson and colleagues (2016) presented participants with short video clips of an actor reaching towards or withdrawing away from an object. Before the onset of the action, participants instructed the actor by saying either "Take it!" or "Leave it!". Participants overestimated the last seen position of the hand in the direction of motion, and this forward bias was larger when the direction of the action followed the instruction. Thus, reaches towards an object appeared more exaggerated after instructing them to "Take it!" than to "Leave it!", and vice versa for observed withdrawals. This Representational Momentum effect was replicated when the actor stated their intention instead, by saying "I'll take it!" or "I'll leave it" before the onset of the action (Hudson et al., 2016), and it was modulated by the likelihood that the actor would do as he says (25% or 75% of the time, Hudson et al, 2018). These findings therefore reveal that high-level information about the goals and intentions of other's actions is

used to generate predictions of their upcoming behaviour and that these predictions create distortions to the perceptual representation of the action, in line the predictive processing theories (Clark, 2013; Friston, 2010).

The experiments in this thesis built on this prior research. They were designed to reveal how these predictions of others' upcoming behaviour are generated, on what cues they rely, and how these predictions alter the perceptual representation of others' action. The following sections will discuss, step by step, the implications of each stage in the action prediction process.

7.2.1.1 Expectations of efficiency

One question addressed by the current work is the prior knowledge that is required to form expectations about others' behaviour. Prior work has revealed a reliance upon social cues such as gaze direction (Frischen, Bayliss, & Tipper, 2007; Hudson, Liu, & Jellema, 2009), facial expression (Hudson & Jellema, 2011b; Yoshikawa & Sato, 2008), and statements of intent (Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016). Here, experiments explored a more general social predictor that has been argued to form the basis of all goal-directed action expectations performed by rational agents: the principle of efficient action (Csibra & Gergely, 2013; Dennett, 1987; Gergely & Csibra, 2003).

Prior research has shown that humans, from a very young age, interpret the actions of others as purposeful and goal-directed (Csibra & Gergely, 2013; Dennett, 1987; Gergely & Csibra, 2003). An essential feature of this "intentional stance" is the realization that humans act rationally, taking the most efficient route possible towards their goals, minimising energy expenditure. Assumptions of efficient action have

been measured in children, as revealed by increased looking times when observed actions violate these assumptions (Gergely & Csibra, 2003; Southgate, Johnson, & Csibra, 2008). In many of these studies, however, action efficiency is measured as a post-hoc evaluation of behaviour that has already been observed. The experiments in this thesis, in contrast, were the first to demonstrate that the principle of efficient action plays a major role in the *formation* of action predictions, forging a perceptual reference action that a rational actor would take to achieve their goal given the environmental constraints (Chapters Three-Five). As such, they do not only establish the principle of action efficiency as a major contributor to goal inference, but also identify the features that motivate predictions of efficiency, both those that are available prior to the action, such as the presence or absence of obstructing objects (Chapter Three) and the agency of the stimuli (Chapter Five), as well as those only revealed during the action, such as its motion profile or the hand's match to target objects (Chapter Six). The principle of action efficiency may therefore act as a general perceptual reference for observed intentional behaviour, structuring action perception in the same way that surrounding illumination structures colour perception (Bloj et al., 1999) and the influence of gravity on motion perception (Hubbard, 1995, 1997), reinforcing the idea that social perception and perception in general may develop from similar neural underpinnings or, at the very least, processing principles.

7.2.1.2 Predictions of action kinematics

Previous studies that utilized a Representational Momentum paradigm to reveal the top-down predictions made during action observation measured perceptual biases towards expectations, such that rotating heads were perceived to have rotated even further when eye-gaze looked towards the same direction (Hudson & Jellema, 2011;

Hudson et al., 2009), and hands were perceived to have reached even closer to goal objects when the actor said “I’ll take it!” compared to “I’ll leave it!” (Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016).

The experiments in this thesis go beyond this prior work to demonstrate that action predictions do not merely reflect abstract goal attribution, creating biases in the general direction of inferred action goals that could be attributed to shifts in attention or anticipatory eye shifts towards these expected goals (Ambrosini, Costantini, & Sinigaglia, 2011). Instead, they show that action predictions reflect the specific kinematics that are required to achieve these goals. The experiments in Chapters Three, Four and Five reveal shifts in perceptual judgements not towards the ultimate goal of the action (the target object) but towards the expected path that the actor should take to reach these objects, governed by the principle of efficient action. Actions that reached inefficiently straight towards an obstacle were perceptually lifted, in line with the expectation of efficiency that the actor would have reached over the obstacle, and reaches that inefficiently arched over an empty space were perceptually lowered, in line with the expectation that a rational actor would have reached straight towards the target. Moreover, making these specific trajectory expectations explicit, as demonstrated in Experiments 1b and 1c, increased perceptual biases. These findings therefore demonstrate that, in line with predictive accounts of perception (Clark, 2013; Friston, 2010), action predictions do not just lead to a bias towards inferred goals. Instead, they provide a perceptual reference trajectory, detailing the specific expected kinematics at each stage of the action, from which the observed action can be compared or integrated, biasing its perceptual representation.

7.2.1.3 Perceptual predictions

Predictive models of non-social perception hypothesise that expectation, derived from prior knowledge and context, may induce changes to even low-level perception (Clark, 2013; Friston, 2010). A famous example is the “dress illusion” whereby inferences about the source of light and surrounding environmental context had a drastic impact on the perceived colour of a dress, so that some people saw it as black and blue while others perceived it as white and gold (Chetverikov & Ivanchei, 2016). Such effects on low-level perception have also been evidenced in the social domain. Apparent motion studies revealed illusory perceptions of motion between two images of a hand that were alternately flashed in two different locations (Chatterjee, Freyd, & Shiffrar, 1996; Shiffrar & Freyd, 1990,1993). The perceived motion did not just reflect lasting after-images on the retina, moved around by shifting eye-movements (Kerzel, 2000), but instead revealed specific expectations about the action kinematics during these gaps, such that the illusory hand movements seem to curve around obstructing objects, revealing the perceptual “filling-in” of the missing information during these intervals (Muckli et al., 2005; Yantis & Nakama, 1998). In the non-social domain, such a filling-in has been confirmed with neuroimaging techniques that reveal corresponding activity in primary visual cortex (Ekman, Kok, & de Lange, 2017; Kok, Brouwer, van Gerven, & de Lange, 2013).

The current thesis adds to prior work by adapting the Representational Momentum paradigm (Hubbard, 2005; Shiffrar & Freyd, 1993), which reliably captures low-level predictive influences on motion perception, with social stimuli (Hudson, Bach, & Nicholson, 2018; Hudson & Jellema, 2011; Hudson, Liu, & Jellema, 2009; Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016).

The current findings do not only extend this work by revealing specific kinematic expectations (see above) but confirm that they act on low-level representations of observed actions, revealing their perceptual nature. Experiment 2a was designed to rule out explanations of these predictive effects from later changes to the action's representation in working memory or motor-related processes during the touch-screen responses. Touch-screen responses were therefore replaced by probe stimuli (a small red dot) that appeared directly after action offset (with 250 ms. interval), either in the predicted direction (e.g. upwards for inefficient straight reaches) or in the unpredicted direction. Probe stimuli in the predicted direction were consistently misjudged to be in the same position as the last seen image of the hand, compared to probes in the unpredicted direction, replicating the predictive effects from experiments in Chapter Three. These findings therefore provide evidence against a working memory or motor-based explanation as prediction effects were apparent immediately after action offset, when working memory does not yet have access to visual representations, and were captured using probe judgements that do not rely on visuospatial motor maps (Kerzel, 2005).

More direct evidence for changes to perception are revealed in Experiment 2b. This study goes one step further than prior work by employing a novel paradigm to verify that these predictive effects do indeed rely on low-level perceptual changes. Experiment 2b utilised a dynamic visual noise mask that reliably disrupts low-level processing and the integration of top-down information on perception (Breitmeyer & Öğmen, 2006; Fahrenfort et al., 2007). The reduction in the perceptual biases confirmed that predictions are perceptually represented and bias the perceptual representation of observed actions.

The experiments in this theses therefore provide evidence that the perceptual representation of action predictions can be similarly explained as other perceptual biases outside social perception. They can therefore either reflect perceptual "filling in" of predicted action kinematics directly after the sudden action offset (Muckli et al., 2005; Yantis & Nakama, 1998) or changes to ongoing perception, reflecting the perceptual sharpening of expected actions (Hammett, 1997; Kok, Jehee, & de Lange, 2012; Yon, Gilbert, Lange, & Press, 2018). Nevertheless, either explanation reflects direct changes to low-level perceptual representations of action towards expected action, providing further support to a predictive account of social perception that corresponds with those underlying non-social perception (Clark, 2013; Friston, 2010).

7.2.1.4 Online action prediction

The current work has revealed that the principle of efficient action is part of the prior knowledge that is required to form assumptions about the goals and intentions of others' behaviour, driving expectations of their forthcoming actions. This adds to prior work that has revealed a reliance upon social cues such as gaze direction (Hudson et al., 2009), facial expression (Hudson & Jellema, 2011b; Yoshikawa & Sato, 2008), and statements of intent (Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016) to generate action expectations. In the experiments in Chapters Three, Four and Five, prior knowledge and contextual information was available from the outset, enabling the formation of specific predictions of upcoming actions before they were initiated. For example, visual access to the target object, the presence or absence of an obstacle, and the agency of the target (hand or ball) was available before the action started, allowing

for the formation of goal assumptions and action predictions well in advance. However, these highly informative contexts do not generalise to the real world where ambiguity and changeability dominate. Often, visual access can be restricted or obscured by overlapping features, by changes in viewpoint orientation or by competition for attention, for example, all of which can change over time (Yuille & Kersten, 2006). The dynamic social world therefore requires an active social prediction system that can be updated online, as more information is revealed. Indeed, the findings from Chapter Five already provide some evidence of a real-time updating of prediction as kinematic information is revealed. In these experiments, the same visual information is available prior to action onset in the Biological and Non-biological ball conditions (i.e. the stationary ball, the target object, and the presence or absence of an obstructing object). Therefore, differences found in predictive effects are likely to be influenced during the ongoing action as more information about from the motion profile is revealed.

The subsequent experiments in this thesis (Chapter 6) directly tested the updating of action predictions when goal information is unavailable prior to action onset and is only revealed dynamically during ongoing action observation. These experiments capitalized on the idea that the affordances of available objects in a scene provide key information about likely goals after the action has already begun (Bach et al., 2014). Specifically, this information can be revealed by the unfolding match between an actor's hand posture and nearby objects such that a small precision grip predicts a reach to a small object and a large power grip to a large object.

Consistent with an online use of such affordance information, the experiments in Chapter Six indeed revealed perceptual biases towards matching target objects. They therefore confirm that such affordance matching information contributes to the

formation of goal assumptions and predictions about how the action will continue. The actor's hand started at rest in every trial, therefore assumptions about which object was the target object was only revealed once the action had initiated and the hand began to form a specific grip type. This therefore provides evidence for an online updating of action prediction during the ongoing action, as grip information is revealed and goal assumptions can be made. The findings go beyond prior work by demonstrating that the goals of others' actions are inferred during the ongoing action, when prior goal information is unavailable, updating action predictions online. They therefore reveal how we so effortlessly navigate the dynamic social world, relying not only on prior knowledge and context, but also on the integration of new information as it is received.

7.2.1.5 Level of representation

The current work has demonstrated that the perceptual representation of observed actions is influenced by the goals that the actor is assumed to hold and how these goals are expected to be achieved. The experiments in Chapter Three show how high-level reasoning about action efficiency guides low-level processing of others' behaviour and the more these efficiency expectations are made explicit, the more they influence action perception. The experiments in Chapter Five and Six reveal that goal assumptions do not only rely on such high-level information that is available prior to action onset, but also on differences in the low-level kinematic features of the action revealed during the ongoing action observation, such as biological motion profiles (Chapter Five) and grip formation (Chapter Six). These findings support prior work from Representational Momentum studies that show that predictions rely on low-level motion features of the stimulus as well as prior knowledge and context

(Hudson, Bach, & Nicholson, 2018; Hudson & Jellema, 2011; Hudson, Liu, & Jellema, 2009; Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016). More broadly, they are consistent with predictive processing models that assume that the perceptual experience of the world emerges from reciprocal interactions between all levels of the cortical hierarchy (Clark, 2013; Csibra, 2008; Friston, 2010; Kilner et al., 2007ab). High-level knowledge cascades down each level of the hierarchy generating a perceptual reference image for sensory input comparisons. This high-level knowledge is constantly updated by new low-level information received by the senses that is transported back up the hierarchy and provides input for the generation of new hypotheses. These models therefore assume that representations on all levels of the hierarchy remain aligned.

Surprisingly, the findings from Chapter Six do not support this alignment of low-level and high-level information. When participants were asked to explicitly judge the goal of the observed action (small strawberry or large apple), although these judgements relied on the same grip-match information as perceptual judgements, they did not correlate on a subject level or on a trial level (Experiment 4a). Those subjects who were more likely to explicitly identify the matched target were not those who showed the largest perceptual biases towards these objects, and the same lack of relationship was revealed on the individual trial level. Moreover, although participants were aware that these goal judgements were required, therefore influencing more explicit processing of the action as it unfolded, this awareness had no effect on the resulting perceptual biases, as revealed by a lack of difference between predictive effects in Experiment 4a compared to Experiment 4b.

A predictive processing account would of course assume a strong relationship between high-level and low-level information, given the connectivity and reciprocity

between each level of the processing hierarchy (Clark, 2013; Friston, 2010). Instead, the results of Chapter Six suggest that, although each level may receive the same input, this information could be processed locally, within the perceptual system for biological motion, for example, as well as globally, across all levels (Firestone & Scholl, 2015; Scholl & Gao, 2013). As a result, the information probed at the perceptual level may contain dissimilarities to that probed at higher levels. Indeed, it has recently been argued that visual illusions show exactly such a distinction, in which high-level knowledge ("I know both lines are of equal length") and perceptual experience ("One line seems longer.") diverge (van Buren & Scholl, 2018).

Indeed, there is now recent evidence that points towards a low-level perceptual evaluation of action intention within the perceptual system, independent of higher-level influence, such as whether one actor chases another (Gao, McCarthy, & Scholl, 2010), whether an actor pays attention to their reach or whether an actor moves certain limbs in a particular direction, dependent on this attention orientation (e.g. moving an arm to the left when the actor is attending to this target position, Jellema, Baker, Wicker, & Perrett, 2000). The results from Chapter Six therefore do not only support such low-level goal attribution, but also suggest that this information is directly translated into a prediction of future action, creating biases to the perceptual representation of the observed action, independent from predictions generated from explicit goal inferences at higher levels.

Such a mechanism would enhance the efficiency of generating rapid predictions of future action, reducing any delay that might result from feedforward and feedback information processing between high-level and low-level regions, for example. By allowing the generation of goal assumptions and associated predictions within the perceptual system itself, the necessary information required to inform own action

selection could be available more quickly, facilitating rapid responsive actions that could be vital for survival.

7.1.2 Understanding the actions of others

Understanding the actions of others is crucial for all social interactions. Through action observation, humans can derive the goals, attitudes and beliefs that drive the actions of others to gain insight into the meaning of their action, and plan a suitable response when required (Bach, Bayliss, & Tipper, 2011; Hamilton, 2009; Sebanz & Knoblich, 2009). Predictive processing theories offer a potential mechanism for how this is accomplished (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007ab). They propose that observers constantly attempt to infer the goals and intentions of others, extracting information from prior knowledge and context to not only guess *what* an actor wants to achieve, but to also predict *how* this will be achieved. Hypotheses about an actor's upcoming actions are then tested by comparing action expectations to observed actions as they unfold, such that a good enough match confirms prior goal assumptions and mismatches are signalled by prediction errors and require a revision to the hypothesis. The experiments in this thesis investigated which information is used to inform these action predictions and what impact these predictions have on social perception, to provide new avenues to understand how action expectations can shape our understanding of other people's behaviour.

The current findings provide evidence for such a top down predictive mechanism of action understanding, revealing the specific kinematic predictions that observers make when watching others' behaviour. They show that these predictions are informed by assumptions of efficient action (Chapter Three), that they are perceptually represented (Chapter Four) and rely on cues to intention from the

target's agency, speed and trajectory (Chapter Five). The results further demonstrate that action predictions do not only rely on prior knowledge and contextual information that is available prior to action onset, but are also dynamically updated during ongoing action observation (Chapter Six). Overall, all experiments in this demonstrate the influence of these action expectations on action perception, distorting their perceptual representation. The findings therefore move away from a purely bottom-up account of action perception (Gazzola & Keysers, 2008; Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010) to support a predictive processing account of social perception (Bach et al., 2014; Csibra, 2008; Kilner, Friston, & Frith, 2007ab), akin to those that underlie perception in general (Clark, 2013; Friston, 2010), and reveal how we so effortlessly navigate the social world.

7.3 Remaining questions and future research

While the findings from the current thesis have provided deeper insights into predictive social perception, revealing the cues that inform predictions of other people's actions and their influence on perception, they have also opened up further questions about the underlying mechanisms. The following sections highlight these questions and propose directions for future research to address them.

7.1.3 Do predictions reflect knowledge of others' minds?

The findings in this thesis confirm that high-level inferences about the goals and intentions of others are translated into perceptual predictions of their upcoming behaviour. In all experiments of this thesis, however, the actor was deemed rational and aware of all target objects and obstacles that might be present, matching the

awareness of the participant. In the real world, other people often possess knowledge and beliefs that can differ from our own, driving their actions in what might be an efficient manner in their reality but not in our own. It therefore remains unclear whether the social predictions captured reflect the observer's own knowledge about the world, or the (potentially false) knowledge attributed to the other person. Moreover, it remains unclear is how these competing cues to intention, from our own reality (e.g. what obstacles we can see) and from assumptions of the reality of others (e.g., what obstacles we think they are aware of), come together to form a prediction of their upcoming action.

Future experiments can resolve this by adapting the efficient actions paradigm used in this thesis. By manipulating whether the actor appears aware of the presence or absence of the obstacle (showing the actor in a blindfold or turned away from the scene, for example), could quantify to what extent perceptual biases reflect the knowledge of the actor or the knowledge of the participant. If social predictions capture the actor's knowledge about the scene, then biases towards more efficient trajectories should be larger when the actor is aware of the scene (i.e. looking towards the objects) compared to when they are not aware (blindfolded or turned away). Such studies could also test whether perceptual predictions capture other's *false* beliefs. By having the actor state their (true or false) beliefs about the scene prior to action onset (e.g. "There is something in the way!" vs. "The path is clear!"), or even changing the scene while they turn away (similar to Sally/Anne tasks, Baron-Cohen, Leslie, & Frith, 1985), these studies could disentangle whether perceptual biases follow the beliefs of the actor or the beliefs of the participant, with the largest shifts expected when these beliefs are aligned.

Such studies would then confirm whether social predictions not only reflect the ambiguity about the content of other's knowledge (e.g. their general awareness of obstacle presence), but also more definitive assumptions about their beliefs that may conflict with reality. Such findings would therefore directly link social predictions to people's inferences about another's mind, as assumed by predictive processing models of social perception (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007ab).

7.1.4 Penetrability of perception

The experiments in Chapters Three to Five have demonstrated how high-level inferences about others' intentions can guide low-level perceptual processing of their actions, with more explicit top-down information exerting more influence on perception. These findings are consistent with predictive processing models that assume that one's perceptual experience of the world emerges from reciprocal interactions between all levels of the cortical hierarchy (Clark, 2013; Csibra, 2008; Friston, 2010; Kilner et al., 2007ab). Strikingly, however, experiments in Chapter Six found a lack of difference between perceptual judgements when goals were explicitly judged and when they were not (Experiment 4a vs Experiment 4b) and a lack of relationship between the perceptual judgements and the explicit goal judgements (Experiment 4a), challenging this strong relationship between high-level and low-level information. Instead, the dissimilarities between information probed at the perceptual and higher levels may reflect that the perceptual effects emerge from locally constrained – or encapsulated – top-down processing within the perceptual system itself, which cannot be penetrated by higher-level knowledge (Firestone & Scholl, 2015; Scholl & Gao, 2013).

One reason why the relationship between high-level knowledge and low-level perception is different in Chapter Six compared to all other experiments could be due to differences in experimental design. First, in Chapter Six, in contrast to all previous experiments, specific goal assumptions could not be made prior to action onset, only unspecific expectations that the actor will reach for one of the two available objects. Any precise high-level inferences can therefore only rely on input from low-level kinematic information once the action has begun. Second, explicit judgements made in Chapter Six were made at the end of the trial and reflect goal identification after the action has been observed and evaluated, rather than goal predictions made at the start of trial before action onset in the previous experiments. These differences suggest that prior knowledge can "tune" the perceptual systems towards expected actions or goals when it is available before action onset, but that these influences may not be possible when such knowledge is only revealed once the action is already underway.

To test whether ongoing action perception is penetrable by high-level knowledge that is only available during the observation, future experiments could utilise the current affordance matching paradigm to manipulate how affordance matching information should be used to reveal action goals, as instructed before each trial. There is an abundance of evidence from behavioural and imaging studies that knowledge of an object's affordance does not only reflect the hand configuration required to move it, but also the potentially different configuration required to use it (Bach, Knoblich, Gunter, Friederici, & Prinz, 2005; Tucker & Ellis, 1998; van Elk, van Schie, & Bekkering, 2014). Moreover, this knowledge is used both to guide action execution (Kalénine, Shapiro, Flumini, Borghi, & Buxbaum, 2014; Valyear, Chapman, Gallivan, Mark, & Culham, 2011) and to predict and understand the actions of others (Bach et

al., 2014, for a review). Future studies could therefore use the same setup as Chapter six, but could inform participants, before each trial, about the actor's intention to either "use" or "move" the object the hand is reaching for. These objects would again be large or small but chosen so that they require different grips to move or use them. For example, a large bottle of fizzy pop requires a large grip to move, but a small grip to unscrew the lid. The differing instructions to "move" or "use" would provide high-level information about the action's general goal, but the action's specific target can only be resolved during the ongoing action, once the hand grip is revealed. If high-level attribution of other's goals can indeed penetrate lower-level perceptual representations, while actions are already underway, then perceptual biases should be recorded towards the object that affords the matching grip type required to achieve the stated goal, not just by the general match between grip size and goal object.

7.1.5 When and how is perception distorted?

The findings from this thesis show that action predictions subtly distort the perceptual representation of observed actions (all chapters), and point to an early perceptual locus of the effects (Chapter Four). The underlying mechanisms for this predictive shaping of social perception are thought to be similar to general mechanisms that allows that sensory input to be interpreted in light of prior expectations (Bach et al., 2014; Csibra, 2008; Kilner, Friston, & Frith, 2007ab). A striking observation is that these top-down influences can act on various stages during stimulus processing. For example, Muckli and colleagues (2005) recorded primary visual cortex (V1) activation during illusory perception in apparent motion tasks, in the precise retinotopic locations for the perceptual "filling-in" of the missing visual information.

Other studies have revealed changes to ongoing perception by sharpening the sensory input from prior expectations (Hammett, 1997; Kok et al., 2012; Yon et al., 2018). Here, perceptual expectations reduced overall V1 activity while improving the precision of the perceptual representation. Perhaps most in line with the current study design is the finding that after training participants with dynamic stimulus sets, only presenting the initial stages of these motions triggers V1 activity that resembles the perception of the full trajectory (Ekman et al., 2017).

In contrast to these prior studies, the perceptual distortions in the current experiments were measured once the action stimuli had disappeared. It is therefore not possible to distinguish whether they appear during ongoing action perception, reflecting perceptual sharpening of their representation, or after action offset, reflecting a perceptual “filling in” of missing information or pre-play of expected action.

These two possibilities can be distinguished with the flash-lag paradigm, designed to probe ongoing motion perception (Kessler, Gordon, Cessford, & Lages, 2010; Nijhawan, 1994, 2008). Such paradigms briefly flash a static stimulus over another moving stimulus. The flash-lag effect marks the well-established finding that the flashed stimulus is perceived to be lagging behind the moving stimulus, even if it is objectively in the same position. It therefore demonstrates a predictive change of the ongoing perceptual representation of the moving target, displaced further along its motion trajectory than it was when the flashed stimulus appeared (Nijhawan, 1994, 2008). If perceptual changes occur already during action observation, then flashing a stimulus over the actor’s hand should elicit a similar lag effect (Kessler et al., 2010). Specifically, for efficient action paradigms, the flashed stimulus should appear to be lower than it actually was when predicting the hand to lift up over an obstacle, and

higher than it actually was when predicting the hand to straighten down towards the target. Similarly, for the affordance matching paradigms, a flashed stimulus should appear to be lower than it actually was when the hand grip matches the top object, and higher when the hand grip matches the bottom object. Such results would reveal an online distortion to the perceptual representation of observed actions towards expectations that can be revealed not only during ongoing action observation, but also after action offset.

Further neuroimaging studies would confirm whether biases to the perceptual representation of observed actions occur during the ongoing observation or in the period after their sudden offset. These studies could utilise modern MVPA methods with fMRI to reveal the patterns of activation in primary visual cortex elicited by predicted actions, to reveal whether action expectations do indeed sharpen their perceptual representation (Ekman et al., 2017; Yon et al., 2018). By training MVPA classifiers to decode specific reach trajectories (straight or arched) from observed actions, these same classifiers should then be able to more easily decode the same actions when they are efficient (and therefore predicted) than when they are inefficient (and therefore unpredicted), as signalled by the presence or absence of an obstacle. Moreover, the neuronal representation of an inefficient straight reach should therefore take on features of the predicted arched reach, and vice versa for inefficient arched reaches. Furthermore, the extent to which inefficient actions elicit activation patterns that more resemble those for an efficient reach should correlate with the size of their subsequent perceptual biases from touch-screen judgements. These findings would confirm a perceptual sharpening account of social perception, revealing ongoing changes to action perception towards expectations.

7.1.6 Do eye-movements play a role in social perception?

The current findings reveal perceptual biases to action representations in the direction of expectations. In the prior literature, such predictive shifts have often been captured by anticipatory eye movements, whereby eye gaze is shifted towards potential goal objects before the actor has reached them (Ambrosini et al., 2011; Costantini, Ambrosini, & Sinigaglia, 2012; Eshuis, Coventry, & Vulchanova, 2009). As noted above, the present results go beyond these findings. In the efficient actions paradigm (Chapters Three, Four and Five), biases were not simply shifted towards the target object, but towards the predicted path that an efficient actor would take to reach this goal (i.e. upwards to avoid an obstacle or straight towards the goal). The results therefore reveal that predictions do not just reflect unspecific goal inferences, but the precise action kinematics that are required to achieve these goals. Nevertheless, since eye movements were not controlled for in these studies, it remains unclear what role they play, if any, in these predictive effects.

The role of predictive eye movements has been hotly debated in the Representational Momentum literature, with inconsistent reports of their influence on perceptual biases (Hubbard, 2005, 2006; Kerzel, 2000, 2003, 2005; Kerzel, Jordan, & Müsseler, 2001). Kerzel (2000) recorded smooth pursuit eye movements during the observation of a continuously moving target and found that they continued to move in the direction of motion for a short period even after target offset. Given that a target persists on the retina for 50-60 milliseconds after its offset, and that forward biases in Representational Momentum disappear when the eyes were fixated, he therefore argued that the representational momentum effect might therefore reflect such ocular overshoots, instead of genuine perceptual changes (Kerzel, 2000).

A problem for such views is that overshoots specifically result from *smooth pursuit* eye movements, as in Kerzel's studies. They therefore do not explain the forward biases elicited by apparent motion and frozen action photographs, when smooth pursuit is not possible (Hubbard, 2005, 2006, 2017), and which therefore mostly likely reflect perceptual prediction. The experiments in this thesis specifically avoided inducing smooth pursuit eye movements for this very reason, by only presenting every third frame of the action, creating non-smooth, "apparent" motion, making a contribution of ocular overshoot due to smooth pursuit eye movements unlikely (Hubbard, 2006, 2017; Kerzel, 2005).

To fully rule out a role of eye movements, future studies could use an eye tracker to capture the ocular movements as perceptual judgements are being made. However, even if eye movements and perceptual biases closely align, this would not provide unequivocal evidence for a role of eye movements in generating perceptual effects. In particular, such findings would not be able to distinguish whether ocular behaviour occurs as a consequence of perceptual change, or whether predictions, instead, steer eye movements towards expected locations, carrying with them visual persistence of the final image (Kerzel, 2000, 2006). In other words, rather than being generated by eye movements, it is possible that perceptual representations have been changed to guide the eyes to predicted future positions (Krauzlis & Stone, 1999). Indeed, such a role has been specifically proposed for representational momentum-like effects (Hubbard, 2006).

To resolve such "cause or effect" questions, future studies could employ variants of the Representational Momentum design in which eye movements are ruled out. For example, eye movements and retinal afterimages are only a problem for movements in the horizontal plane (i.e. actions in side view), but not for movements in depth (i.e.

towards/away from oneself), which primarily lead to objects becoming visually larger or smaller (Webster, 2011). Consider a task in which participants see a static image of an actor holding an object out to them. The actor then either says “I’ll keep it for me!” or “I’ll give it to you!” to elicit expectations of forwards or backwards movements. The hand would then either reach out towards the participant, increasing in size as it nears, or withdraw away from the participant and reduce in size. The sudden offset of the action would be replaced with a probe image that is either slightly larger (as if it had moved closer), smaller (as if it was further away) or the same size as the last image. If goal attribution directly shapes action perception, then approaching actions should appear even larger when an approach was expected, compared to a withdrawal and withdrawing actions should appear even smaller when a withdrawal was expected, compared to an approach, and these perceptual biases would be reflected in their probe judgements. Such results would provide direct evidence that perceptual changes can be induced in the absence of eye movements (Kerzel, 2000, 2006), cannot be explained by visual persistence (the image changes in size), and that they therefore result from changes to the action’s perceptual representation.

7.1.7 How are perceptual predictions updated?

The findings in this thesis support predictive processing theories of social perception (Bach et al., 2014; Csibra, 2008; Kilner, Friston, & Frith, 2007ab) by revealing the top-down influence of expectation on action observation that distort their perceptual representation. These predictive processing models view perception as a process of hypothesis testing *and revision*. However, the experiments so far only tested the downstream effect of prior knowledge, as measured by its effects on perception.

They leave open whether large enough mismatches between predictions and observations, in turn, are propagated back upwards to change prior inferences of the action. What therefore remains to be determined is the extent to which goal assumptions are revised when the incoming information does not support the initial hypothesis, as signalled by prediction errors (e.g. when clearly seeing a straight reach despite an obstacle being in the way).

In predictive processing models (e.g. Clark, 2013; Friston, 2010), perceptual representations emerge from a Bayesian integration of what was predicted and what was observed, with more weight being placed on expectations the more the input is ambiguous (Kok et al., 2013; Lages, Boyle, & Jenkins, 2017). The perceptual biases captured here therefore reflect the influence of top down predictions when observing short and rapid action clips, which are subject to considerable uncertainty (Kok et al., 2013; Lages et al., 2017). Such settings are typical in the real world, where multiple sources of information are combined to generate expectations of others' behaviour which, when observed, occur rapidly, often behind multiple occluders and in environments with endless opportunities for action. For example, at a busy road crossing, you just know that the hurrying pedestrian with their hand held out will press the button, perhaps multiple times, despite many cars and other pedestrians obscuring your view of their action. Predictions, in such models, are therefore more likely to be updated the more precise (less ambiguous) the input becomes, and the more it diverges from prior expectations (Clark, 2013; Friston & Kiebel, 2009).

The updating of goal assumptions can be directly tested by introducing post-action judgements at the end of each trial, similar to the experiments in Chapter Six. In the context of the action efficiency experiments in this thesis, for example, participants would still report the hand's last seen position (via touch or probe judgments), but

additionally report how confident they are about what they saw. These judgements could probe both the kinematics features that participants believe they saw (i.e. did the actor make a straight or arched reach?) as well as high-level judgements about inferred mental states (i.e. did the actor see the obstacle or not?). By manipulating the ambiguity of the input – for example, by varying how long each action is seen – revisions of prior expectations can be tracked as the input becomes more precise. One would expect that these explicit judgments follow the initial expectations provided by the context (e.g. an obstacle in the way) the more ambiguous the action input is but diverge from these expectations – and become updated – the more precise it is. Importantly, one could then again track the relationship between explicit judgments and perceptual biases. In particular, it is expected that both measures would be closely related across trials, such that the more an inefficient reach towards as obstacle is perceptually lifted over the obstacle, the less likely they are to detect that the reach was actually straight, for example. Such findings would point to a close alignment of perceptual and high-level judgments, as assumed by predictive processing models of social perception (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007ab).

The results from the current experiments (Chapters Three, Four and Five) already point to such a process. They imply that goal assumptions are not revised, and that participants continue to believe that the actor is aware of the presence or absence of obstacles (otherwise they would not show perceptual biases). They suggest that perceptual biases towards expectations may act as a “perceptual confirmation bias” that prevents this updating, limiting the detection of a mismatch.

Future experiments could track this interplay between perceptual evidence and prior expectation and reveal at what point the mismatch become too large, such that

perception is no longer biased, and a revision to the prediction is forced instead. To resolve this question, future studies can utilise the current paradigm and subtly manipulate the kinematics of the action, such that the observed action is gradually shifted away from the predicted trajectory. For example, some actions would clearly resemble a straight reach, whereas others vary from a slight arch to a more extreme arched trajectory. It is expected that, initially, a perceptual bias would be induced, so that a slight arch remains undetected. As the action becomes progressively further from the predicted straight reach, the point at which a difference is perceived can be captured by gradual changes in perceptual biases and post-action judgements.

7.4 Summary and conclusion

Understanding the actions of others is crucial for all social interactions. Despite a dynamic and complicated social world, humans can derive the goals, attitudes and beliefs that drive the actions of others, imbuing them with meaning and understanding (Hamilton & Grafton, 2007; Press, Heyes, & Kilner, 2011). The current thesis revealed that predictions play a large part in these abilities. By integrating the knowledge about others (e.g. their mental states) with observations of their actions, expectations of their upcoming movements can be derived (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007ab). These action predictions can then provide a perceptual reference image that guides observation, akin to processes in non-social perception (Clark, 2013; Friston, 2010; Friston & Kiebel, 2009). The current research showed (1) that such expectations take place during social perception, (2) that they shape the observation of others' action, (3) in line with the principle of efficient action, (4) guided by the goals and intentions attributed to the other person. It therefore confirms that the perceptual representation of others' actions is biased by

the intentions we assign to them and our expectations of how these intentions will be fulfilled. Future research must now address how multiple sources of intention information are integrated, how they are translated into low-level kinematic predictions, when and how they influence perception, and how they are updated by new information.

7.2 References

Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neurophysiology*, *106*(3), 1437–1442. https://doi.org/10.1152/jn.00118.2011

Aziz-Zadeh, L., Koski, L., Zaidel, E., Mazziotta, J., & Iacoboni, M. (2006). Lateralization of the human mirror neuron system. *Journal of Neuroscience*, *26*(11), 2964–2970.

Aziz-Zadeh, L., Maeda, F., Zaidel, E., Mazziotta, J., & Iacoboni, M. (2002). Lateralization in motor facilitation during action observation: a TMS study. *Experimental Brain Research*, *144*(1), 127–131.

Bach, P., Bayliss, A. P., & Tipper, S. P. (2010). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychonomic Bulletin & Review*, *18*(1), 171–176. https://doi.org/10.3758/s13423-010-0029-x

Bach, P., Bayliss, A. P., & Tipper, S. P. (2011). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychonomic Bulletin & Review*, *18*(1), 171–176. https://doi.org/10.3758/s13423-010-0029-x

Bach, P., Knoblich, G., Gunter, T. C., Friederici, A. D., & Prinz, W. (2005a). Action Comprehension: Deriving Spatial and Functional Relations. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 465–479. https://doi.org/10.1037/0096-1523.31.3.465

Bach, P., Knoblich, G., Gunter, T. C., Friederici, A. D., & Prinz, W. (2005b). Action comprehension: deriving spatial and functional relations. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 465.

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00254

Bach, P., Peatfield, N. A., & Tipper, S. P. (2007). Focusing on body sites: the role of spatial attention in action perception. *Experimental Brain Research*, *178*(4), 509–517.

Bach, P., & Schenke, K. C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass*, *11*(7), e12312. https://doi.org/10.1111/spc3.12312

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 210.

Bex, P. J., Edgar, G. K., & Smith, A. T. (1995). Sharpening of drifting, blurred images. *Vision Research*, *35*(18), 2539–2546. https://doi.org/10.1016/0042-6989(95)00060-d

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, *402*(6764), 877–879. https://doi.org/10.1038/47245

Bozzacchi, C., Spinelli, D., Pitzalis, S., Giusti, M. A., & Di Russo, F. (2015). I know what I will see: action-specific motor preparation activity in a passive observation task. *Social Cognitive and Affective Neuroscience*, *10*(6), 783–789. https://doi.org/10.1093/scan/nsu115

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1–2), 3–22.

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495.

Breitmeyer, B., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision*. Oxford University Press.

Buxbaum, L. J., Johnson-Frey, S. H., & Bartlett-Williams, M. (2005). Deficient internal models for planning hand–object interactions in apraxia. *Neuropsychologia*, *43*(6), 917–929.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2004). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral Cortex*, *15*(8), 1243–1249.

Chetverikov, A., & Ivanchei, I. (2016). Seeing “the Dress” in the Right Light: Perceived Colors and Inferred Light Sources. *Perception*, *45*(8), 910–930. https://doi.org/10.1177/0301006616643664

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(03), 181–204. https://doi.org/10.1017/s0140525x12000477

Costantini, M., Ambrosini, E., & Sinigaglia, C. (2012). Out of your hand’s reach, out of my eyes’ reach. *Quarterly Journal of Experimental Psychology*, *65*(5), 848–855. https://doi.org/10.1080/17470218.2012.679945

Cross, E. S., Hamilton, A. F. de C., & Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *Neuroimage*, *31*(3), 1257–1267.

Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorymotor foundations of higher cognition. Attention and performance XXII* (pp. 435–459). Oxford: Oxford University Press.

Csibra, G., & Gergely, G. (2013). Teleological understanding of actions. *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*, 38–43.

D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, *24*(3), 955–958.

de Lafuente, V., & Romo, R. (2004). Language abilities of motor cortex. *Neuron*, *41*(2), 178–180.

Dennett, D. C. (1987). *The intentional stance.* Cambridge, MA: MIT Press.

Desmet, C., & Brass, M. (2015). Observing accidental and intentional unusual actions is associated with different subregions of the medial frontal cortex. *NeuroImage*, *122*, 195–202.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*(1), 176–180.

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*, 15276. https://doi.org/10.1038/ncomms15276

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009a). Predictive Eye Movements Are Driven by Goals, Not by the Mirror Neuron System. *Psychological Science*, *20*(4), 438–440. https://doi.org/10.1111/j.1467-9280.2009.02317.x

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009b). Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychological Science*, *20*(4), 438–440.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*(6), 2608–2611.

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2007). Masking Disrupts Reentrant Processing in Human Visual Cortex. *Journal of Cognitive Neuroscience*, *19*(9), 1488–1497. https://doi.org/10.1162/jocn.2007.19.9.1488

Finke, R. A., Freyd, J. J., & Shyi, G. C. (1986). Implied velocity and acceleration induce transformations of visual memory. *Journal of Experimental Psychology: General*, *115*(2), 175.

Firestone, C., & Scholl, B. J. (2015). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/s0140525x15000965

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, *308*(5722), 662–667.

Fontana, A. P., Kilner, J. M., Rodrigues, E. C., Joffily, M., Nighoghossian, N., Vargas, C. D., & Sirigu, A. (2012). Role of the parietal cortex in predicting incoming actions. *Neuroergonomics: The Human Brain in Action and at Work*, *59*(1), 556–564. https://doi.org/10.1016/j.neuroimage.2011.07.046

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, *33*(6), 575–581.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126–132. https://doi.org/10.1037//0278-7393.10.1.126

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society*, *23*(6), 443–446. https://doi.org/10.3758/BF03329847

Freyd, J. J., & Jones, K. T. (1994). Representational momentum for a spiral path. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 968.

Freyd, J. J., & Pantzer, T. M. (1995). *Static patterns moving in the mind.*

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127.

Friston, K., & Kiebel, S. (2009a). Cortical circuits for perceptual inference. *Neural Networks*, *22*(8), 1093–1104. https://doi.org/10.1016/j.neunet.2009.07.023

Friston, K., & Kiebel, S. (2009b). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996a). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996b). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). *17 Action representation and the inferior parietal lobule*.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501.

Gallese, V., & Sinigaglia, C. (2010). The bodily self as power for action. *Neuropsychologia*, *48*(3), 746–755.

Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2004). Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. *European Journal of Neuroscience*, *20*(8), 2193–2202.

Gao, T., McCarthy, G., & Scholl, B. J. (2010). The Wolfpack Effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science*, *21*(12), 1845–1853. https://doi.org/10.1177/0956797610388814

Gazzola, V., & Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, *19*(6), 1239–1255.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1

Gillmeister, H., Catmur, C., Liepelt, R., Brass, M., & Heyes, C. (2008). Experience-based priming of body parts: A study of action imitation. *Brain Research*, *1217*, 157–170. https://doi.org/10.1016/j.brainres.2007.12.076

Grèzes, J. (1998). Top down effect of strategy on the perception of human biological motion: A PET investigation. *Cognitive Neuropsychology*, *15*(6–8), 553–582.

Grèzes, J., Costes, N., & Decety, J. (1999). The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain*, *122*(10), 1875–1887.

Hamilton, A. (2009). Research review: Goals, intentions and mental states: challenges for theories of autism. *Journal of Child Psychology and Psychiatry*, *50*(8), 881–892. https://doi.org/10.1111/j.1469-7610.2009.02098.x

Hamilton, A., & Grafton, S. T. (2007). The motor hierarchy: from kinematics to goals and intentions. *Sensorimotor Foundations of Higher Cognition*, *22*, 381–408.

Hammett, S. T. (1997). Motion blur and motion sharpening in the human visual system. *Vision Research*, *37*(18), 2505–2510. https://doi.org/10.1016/s0042-6989(97)00059-x

Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, *17*(2), 282–293.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307.

Heptulla Chatterjee, S., Freyd, J. J., & Shiffrar, M. (1996). Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 916.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, *21*(7), 1229–1243.

Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition*, *108*(3), 687–701.

Hubbard, T. L. (1995a). Cognitive representation of motion: Evidence for friction and gravity analogues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 241.

Hubbard, T. L. (1995b). Cognitive representation of motion: Evidence for friction and gravity analogues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 241–254. https://doi.org/10.1037//0278-7393.21.1.241

Hubbard, T. L. (1996). Representational momentum, centripetal force, and curvilinear impetus. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(4), 1049.

Hubbard, T. L. (1997a). Target size and displacement along the axis of implied gravitational attraction: Effects of implied weight and evidence of representational gravity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(6), 1484.

Hubbard, T. L. (1997b). Target size and displacement along the axis of implied gravitational attraction: Effects of implied weight and evidence of representational gravity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(6), 1484.

Hubbard, T. L. (1997c). Target size and displacement along the axis of implied gravitational attraction: Effects of implied weight and evidence of representational gravity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(6), 1484.

Hubbard, T. L. (1998). Some effects of representational friction, target size, and memory averaging on memory for vertically moving targets. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *52*(1), 44.

Hubbard, T. L. (2001). The effect of height in the picture plane on the forward displacement of ascending and descending targets. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *55*(4), 325.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, *12*(5), 822–851. https://doi.org/10.3758/bf03196775

Hubbard, T. L. (2006). Computational theory and cognition in representational momentum and related types of displacement: A reply to Kerzel. *Psychonomic Bulletin & Review*, *13*(1), 174–177.

Hubbard, T. L. (2015). The varieties of momentum-like experience. *Psychological Bulletin*, *141*(6), 1081–1119. https://doi.org/10.1037/bul0000016

Hubbard, T. L. (2017). Toward a general theory of momentum-like effects. *Behavioural Processes*, *141*, 50–66.

Hubbard, T. L., & Bharucha, J. J. (1988a). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, *44*(3), 211–221.

Hubbard, T. L., & Bharucha, J. J. (1988b). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, *44*(3), 211–221.

Hubbard, T. L., & Favretto, A. (2003). Naïve impetus and Michotte's 'tool effect': evidence from representational momentum. *Psychological Research*, *67*(2), 134–152. https://doi.org/10.1007/s00426-002-0122-5

Hubbard, T. L., & Ruppel, S. E. (2002). A possible role of naïve impetus in Michotte's 'launching effect': Evidence from representational momentum. *Visual Cognition*, *9*(1–2), 153–176. https://doi.org/10.1080/13506280143000377

Hudson, M., Bach, P., & Nicholson, T. (2018). You said you would! The predictability of other's behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(2), 320–335. https://doi.org/10.1037/xhp0000451

Hudson, Matthew, Burnett, H. G., & Jellema, T. (2012). Anticipation of action intentions in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *42*(8), 1684–1693.

Hudson, Matthew, & Jellema, T. (2011a). Resolving ambiguous behavioral intentions by means of involuntary prioritization of gaze processing. *Emotion*, *11*(3), 681–686. https://doi.org/10.1037/a0023264

Hudson, Matthew, & Jellema, T. (2011b). Resolving ambiguous behavioral intentions by means of involuntary prioritization of gaze processing. *Emotion*, *11*(3), 681–686. https://doi.org/10.1037/a0023264

Hudson, Matthew, Liu, C. H., & Jellema, T. (2009a). Anticipating intentional actions: The effect of eye gaze direction on the judgment of head rotation. *Cognition*, *112*(3), 423–434. https://doi.org/10.1016/j.cognition.2009.06.011

Hudson, Matthew, Liu, C. H., & Jellema, T. (2009b). Anticipating intentional actions: The effect of eye gaze direction on the judgment of head rotation. *Cognition*, *112*(3), 423–434. https://doi.org/10.1016/j.cognition.2009.06.011

Hudson, Matthew, Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition*, *146*, 245–250. https://doi.org/10.1016/j.cognition.2015.09.021

Hudson, Matthew, Nicholson, T., Simpson, W. A., Ellis, R., & Bach, P. (2016). One step ahead: The perceived kinematics of others' actions are biased toward expected goals. *Journal of Experimental Psychology: General*, *145*(1), 1–7. https://doi.org/10.1037/xge0000126

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, *60*, 653–670.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the Intentions of Others with One's Own Mirror Neuron System. *PLOS Biology*, *3*(3), e79. https://doi.org/10.1371/journal.pbio.0030079

Ianì, F., Mazzoni, G., & Bucciarelli, M. (2018). The role of kinematic mental simulation in creating false memories. *Journal of Cognitive Psychology*, *30*(3), 292–306.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, *9*(1), 21–25.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, *14*(1), S103–S109.

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural Representation for the Perception of the Intentionality of Actions. *Brain and Cognition*, *44*(2), 280–302. https://doi.org/10.1006/brcg.2000.1231

Johnston, H. M., & Jones, M. R. (2006). Higher order pattern structure influences auditory representational momentum. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 2.

Jonas, M., Siebner, H., Biermann-Ruben, K., Kessler, K., Bäumer, T., Büchel, C., … Münchau, A. (2007). Do simple intransitive finger movements consistently activate frontoparietal mirror neuron areas in humans? *Neuroimage*, *36*, T44–T53.

Jordan, J. S., & Hunsinger, M. (2008). Learned patterns of action-effect anticipation contribute to the spatial displacement of continuously moving stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 113–124. https://doi.org/10.1037/0096-1523.34.1.113

Kalénine, S., Shapiro, A. D., Flumini, A., Borghi, A. M., & Buxbaum, L. J. (2014). Visual context modulates potentiation of grasp types during semantic object categorization. *Psychonomic Bulletin & Review*, *21*(3), 645–651.

Kerzel, D. (2000). Eye movements and visible persistence explain the mislocalization of the final position of a moving target. *Vision Research*, *40*(27), 3703–3715. https://doi.org/10.1016/S0042-6989(00)00226-1

Kerzel, D. (2003). Mental extrapolation of target position is strongest with weak motion signals and motor responses. *Vision Research*, *43*(25), 2623–2635.

Kerzel, D. (2005). Representational momentum beyond internalized physics: Embodied mechanisms of anticipation cause errors in visual short-term memory. *Current Directions in Psychological Science*, *14*(4), 180–184.

Kerzel, D., Jordan, J. S., & Müsseler, J. (2001). The role of perception in the mislocalization of the final position of a moving target. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 829–840. https://doi.org/10.1037/0096-1523.27.4.829

Kessler, K., Gordon, L., Cessford, K., & Lages, M. (2010). Characteristics of motor resonance predict the pattern of flash-lag effects for biological motion. *PloS One*, *5*(1), e8258.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, *153*(4), 628–636.

Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences*, *8*(11), 501–507.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007a). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007b). The mirror-neuron system: a Bayesian perspective. *NeuroReport*, *18*(6), 619–623. https://doi.org/10.1097/wnr.0b013e3281139ed0

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, *7*(12), 1299.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, *297*(5582), 846–848.

Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience*, *33*(41), 16275–16284.

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, *75*(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kolers, P. A. (2013). *Aspects of Motion Perception: International Series of Monographs in Experimental Psychology* (Vol. 16). Elsevier.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, *12*(1), 48–55.

Krauzlis, R. J., & Stone, L. S. (1999). Tracking with the mind's eye. *Trends in Neurosciences*, *22*(12), 544–550.

Lages, M., Boyle, S. C., & Jenkins, R. (2017). Illusory Increases in Font Size Improve Letter Recognition. *Psychological Science*, *28*(8), 1180–1188. https://doi.org/10.1177/0956797617705391

Liepelt, R., Von Cramon, D. Y., & Brass, M. (2008). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *Neuroimage*, *43*(4), 784–792.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1–3), 59–70.

Maranesi, M., Livi, A., Fogassi, L., Rizzolatti, G., & Bonini, L. (2014). Mirror neuron activation prior to action observation in a predictable context. *Journal of Neuroscience*, *34*(45), 14827–14832.

Marsh, L. E., Mullett, T. L., Ropar, D., & Hamilton, A. F. de C. (2014). Responses to irrational actions in action observation and mentalising networks of the human brain. *NeuroImage*, *103*, 81–90.

MCGURK, H., & MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary Visual Cortex Activity along the Apparent-Motion Trace Reflects Illusory Perception. *PLoS Biology*, *3*(8), e265. https://doi.org/10.1371/journal.pbio.0030265

Nagai, M., Kazai, K., & Yagi, A. (2002). Larger forward memory displacement in the direction of gravity. *Visual Cognition*, *9*(1–2), 28–40.

Negri, G. A., Rumiati, R. I., Zadini, A., Ukmar, M., Mahon, B. Z., & Caramazza, A. (2007). What is the role of motor simulation in action and object recognition? Evidence from apraxia. *Cognitive Neuropsychology*, *24*(8), 795–816.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*(6487), 256–257. https://doi.org/10.1038/370256b0

Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences*, *31*(02). https://doi.org/10.1017/s0140525x08003804

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Viewpoint (in) dependence of action representations: an MVPA study. *Journal of Cognitive Neuroscience*, *24*(4), 975–989.

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2013). Crossmodal and action-specific: neuroimaging the human mirror neuron system. *Trends in Cognitive Sciences*, *17*(7), 311–318.

Pazzaglia, M., Pizzamiglio, L., Pes, E., & Aglioti, S. M. (2008). The sound of actions in apraxia. *Current Biology*, *18*(22), 1766–1772.

Press, C., Heyes, C., & Kilner, J. M. (2011). Learning to understand others' actions. *Biology Letters*, *7*(3), 457–460. https://doi.org/10.1098/rsbl.2010.0850

Press, Clare, Bird, G., Walsh, E., & Heyes, C. (2008). Automatic imitation of intransitive actions. *Brain and Cognition*, *67*(1), 44–50. https://doi.org/10.1016/j.bandc.2007.11.001

Reed, C. L., & Vinson, N. G. (1996). Conceptual effects on representational momentum. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 839.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169–192.

Rizzolatti, G., & Fabbri-Destro, M. (2008). The mirror system and its role in social cognition. *Current Opinion in Neurobiology*, *18*(2), 179–184.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*(4), 264–274. https://doi.org/10.1038/nrn2805

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, *119*(2), 242–252. https://doi.org/10.1016/j.cognition.2011.01.014

Schenke, K. C., Wyer, N. A., & Bach, P. (2016). The Things You Do: Internal Models of Others' Expected Behaviour Guide Action Observation. *PLOS ONE*, *11*(7), e0158910. https://doi.org/10.1371/journal.pone.0158910

Scholl, B. J., & Gao, T. (2013). Perceiving Animacy and Intentionality: Visual processing or higher-level judgment. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social Perception: Detection and interpretation of animacy, agency, and intention,* (pp. 197–230). https://doi.org/10.7551/mitpress/9780262019279.003.0009

Sebanz, N., & Knoblich, G. (2009). Prediction in Joint Action: What, When, and Where. *Topics in Cognitive Science*, *1*(2), 353–367. https://doi.org/10.1111/j.1756-8765.2009.01024.x

Senior, C., Barnes, J., Giampietroc, V., Simmons, A., Bullmore, E., Brammer, M., & David, A. (2000a). The functional neuroanatomy of implicit-motion perception or 'representational momentum'. *Current Biology*, *10*(1), 16–22.

Senior, C., Barnes, J., Giampietroc, V., Simmons, A., Bullmore, E., Brammer, M., & David, A. (2000b). The functional neuroanatomy of implicit-motion perception or 'representational momentum'. *Current Biology*, *10*(1), 16–22.

Senior, C., Ward, J., & David, A. S. (2002). Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Visual Cognition*, *9*(1–2), 81–92.

Shiffrar, M., & Freyd, J. J. (1993). Timing and Apparent Motion Path Choice With Human Body Photographs. *Psychological Science*, *4*(6), 379–384. https://doi.org/10.1111/j.1467-9280.1993.tb00585.x

Shmuelof, L., & Zohary, E. (2006). A mirror representation of others' actions in the human anterior parietal cortex. *Journal of Neuroscience*, *26*(38), 9736–9742.

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition*, *107*(3), 1059–1069.

Southgate, V., Johnson, M. H., Karoui, I. E., & Csibra, G. (2010). Motor System Activation Reveals Infants' On-Line Prediction of Others' Goals. *Psychological Science*, *21*(3), 355–359. https://doi.org/10.1177/0956797610362058

Southgate Victoria, Johnson Mark H., Osborne Tamsin, & Csibra Gergely. (2009). Predictive motor activation during action observation in human infants. *Biology Letters*, *5*(6), 769–772. https://doi.org/10.1098/rsbl.2009.0474

Sterzer, P., Haynes, J.-D., & Rees, G. (2006). Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT+/V5. *Neuroimage*, *32*(3), 1308–1316.

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1746.

Thornton, I. M., & Hubbard, T. L. (2002). Representational momentum: New findings, new directions. *Visual Cognition*, *9*(1–2), 1–7.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 830.

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001a). I know what you are doing: A neurophysiological study. *Neuron*, *31*(1), 155–165.

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001b). I know what you are doing: A neurophysiological study. *Neuron*, *31*(1), 155–165.

Valyear, K. F., Chapman, C. S., Gallivan, J. P., Mark, R. S., & Culham, J. C. (2011). To use or to move: goal-set modulates priming when grasping real tools. *Experimental Brain Research*, *212*(1), 125–142.

van Elk, M., van Schie, H., & Bekkering, H. (2014). Action semantics: A unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Physics of Life Reviews*, *11*(2), 220–250. https://doi.org/10.1016/j.plrev.2013.11.005

Verfaillie, K., & d'Ydewalle, G. (1991). Representational momentum and event course anticipation in the perception of implied periodical motions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(2), 302.

Vinson, N. G., & Reed, C. L. (2002). Sources of object-specific effects in representational momentum. *Visual Cognition*, *9*(1–2), 41–65.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5), 3–3.

Weiss, P. H., Rahbari, N. N., Hesse, M. D., & Fink, G. R. (2008). Deficient sequencing of pantomimes in apraxia. *Neurology*, *70*(11), 834–840.

Wellman, H. M., & Brandone, A. C. (2009). Early intention understandings that are common to primates predict children's later theory of mind. *Current Opinion in Neurobiology*, *19*(1), 57–62. https://doi.org/10.1016/j.conb.2009.02.004

Wertheimer, M. (1912). Experimentelle studien uber das sehen von bewegung. *Zeitschrift Fur Psychologie*, *61*.

Wiggett, A. J., Hudson, M., Tipper, S. P., & Downing, P. E. (2011). Learning associations between action and perception: effects of incompatible training on body part and spatial priming. *Brain and Cognition*, *76*(1), 87–96.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460.

Windmann, S. (2004). Effects of sentence context and expectation on the McGurk illusion. *Journal of Memory and Language*, *50*(2), 212–230.

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience*, *1*(6), 508–512. https://doi.org/10.1038/2226

Yon, D., Gilbert, S. J., Lange, F. P. de, & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, *9*(1), 4288. https://doi.org/10.1038/s41467-018-06752-7

Yoshikawa, S., & Sato, W. (2008). Dynamic facial expressions of emotion induce representational momentum. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 25–31. https://doi.org/10.3758/CABN.8.1.25

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301–308.